

(19)日本国特許庁 (JP)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開2002-91991

(P2002-91991A)

(43)公開日 平成14年3月29日 (2002.3.29)

(51)Int.Cl.⁷

G 0 6 F 17/30

識別記号

1 7 0

F I

デコード(参考)

3 6 0

G 0 6 F 17/30

1 7 0 F 4 B 0 2 4

4 1 9

3 6 0 Z 4 B 0 2 9

C 1 2 M 1/00

4 1 9 B 5 B 0 7 5

C 1 2 N 15/09

C 1 2 M 1/00

A

C 1 2 N 15/00

A

審査請求 未請求 請求項の数17 O.L. (全33頁)

(21)出願番号

特願2000-284612(P2000-284612)

(71)出願人 500309920

インテック・ウェブ・アンド・ゲノム・インフォマティクス株式会社

東京都江東区新砂1-3-3

(22)出願日

平成12年9月20日 (2000.9.20)

(72)発明者 吉田 美寸夫

東京都江東区新砂一丁目3番3号 インテック・ウェブ・アンド・ゲノム・インフォマティクス株式会社内

(74)代理人 10007/540

弁理士 高野 昌俊

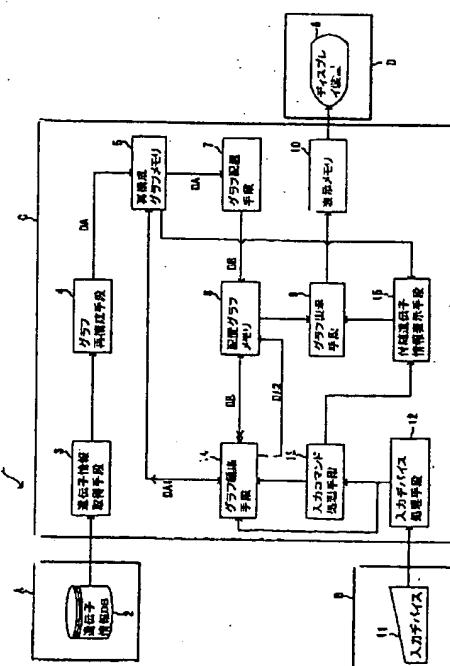
最終頁に続く

(54)【発明の名称】 遺伝子ネットワーク研究支援システム及び方法

(57)【要約】

【課題】 遺伝子に関する多種、多様なデータに基づいて遺伝子ネットワークを構築するのを支援すること。

【解決手段】 遺伝子情報データベース2に蓄積した要素情報及び関係情報を用いて遺伝子ネットワークを再構成し、ディスプレイ装置6上にネットワーク図を表示しておき、ユーザが入力デバイス11から編集コマンドを入力することにより、グラフ編集手段14によって、表示されているネットワーク図の点または辺を変更、削除、追加し、あるいは付随情報を表示させる等の遺伝子ネットワークの編集を行うことができるようにした。



BEST AVAILABLE COPY

(2) 開2002-91991 (P2002-91991A)

【特許請求の範囲】

【請求項1】 コンピュータを用い遺伝子に関する情報を統合して遺伝子ネットワークとして視覚的に表示することにより遺伝子研究を支援するシステムであって、遺伝子、RNA、タンパク質、タンパク質複合体、低分子、環境条件、細胞周期、細胞タイプ、表現型などの要素情報と要素間の結合や制御などの関係情報を蓄積したデータベースと、データベースに蓄積される要素情報をノード、関係情報をエッジとして表現される連結成分を生成するグラフ再構成手段と、

前記連結成分をグラフィカル表示する表示手段と、前記表示手段によって表示されているグラフのトポロジーを変更するグラフ編集手段とを備えたことを特徴とする遺伝子ネットワーク研究支援システム。

【請求項2】 前記グラフ再構成手段によって生成された複数の連結成分から特定の連結成分だけを選択的に表示できるようにした請求項1記載の遺伝子ネットワーク研究支援システム。

【請求項3】 前記グラフ再構成手段に対して、連結成分を生成する過程で関係同士を結び付けるか否かの条件・制約を指定することができるようとした請求項1記載の遺伝子ネットワーク研究支援システム。

【請求項4】 既存の文献あるいはデータベースから情報を抽出する情報抽出手段と、該情報抽出手段によって得られた情報を前記データベースに蓄積する情報保存手段とをさらに含む請求項1記載の遺伝子ネットワーク研究支援システム。

【請求項5】 遺伝子情報を統合してグラフとして視覚的に表示する遺伝子ネットワーク研究支援方法であって、

(a) 関係情報を追加・削除するユーザ入力を受け取るステップと、

(b) 追加・削除された関係情報を加味してグラフを再構成するステップと、

(c) 再構成されたグラフを視覚的に表示するステップとを備えたことを特徴とする遺伝子ネットワーク研究支援方法。

【請求項6】 前記ステップ(a)から(c)が複数回反復される請求項5記載の遺伝子ネットワーク研究支援方法。

【請求項7】 遺伝子情報を統合してグラフとして視覚的に表示する遺伝子ネットワーク研究支援方法であって、

(a) ノードあるいはエッジを指定するためのユーザ入力を受け取るステップと、

(b) 指定されたノードあるいはエッジに対応する要素に付随する情報を前記データベースから取得するステップと、

(c) 取得された付随情報を前記グラフに関連させて表

示するステップとを備えたことを特徴とする遺伝子ネットワーク研究支援方法。

【請求項8】 前記付随情報が、タンパク質局在性、発現プロファイル、タンパク質機能、タンパク質機能ドメイン情報を有している請求項7記載の遺伝子ネットワーク研究支援方法。

【請求項9】 付随する情報に応じて、点・辺に対応する表示プリミティブの視覚的特徴を変更する手段を備えた請求項1記載の遺伝子ネットワーク研究支援システム。

【請求項10】 演繹推論により、要素間の関係に介在する要素、関係を取り除く要素間関係推論を行う演繹推論部を備えた請求項1記載の遺伝子ネットワーク研究支援システム。

【請求項11】 前記要素間関係推論の結果が表示されている場合に、要素間関係に対応するエッジを指定することにより、指定されたエッジに対応する要素間関係に介在する関係情報を対応するエッジを表示するようにした請求項10記載の遺伝子ネットワーク研究支援システム。

【請求項12】 前記データベースに蓄積された要素情報および関係情報を付随する情報に応じて、それらに対応する点・辺の色の濃度を変更するようにした請求項9記載の遺伝子ネットワーク研究支援システム。

【請求項13】 複数のグラフを重ね合わせてレイヤー表示するレイヤー表示手段を備えた請求項1記載の遺伝子ネットワーク研究支援システム。

【請求項14】 前記レイヤー表示手段によって表示される、複数のレイヤー間で重複する又は重複しない点・辺を強調表示するようにした請求項13記載の遺伝子ネットワーク研究支援方法。

【請求項15】 遺伝子、タンパク質が出現する文献情報を取得するステップと、関係を構成する遺伝子、タンパク質間で重複する文献情報を求めるステップと、得られた結果を表示するステップとを備えたことを特徴とする遺伝子ネットワーク研究支援方法。

【請求項16】 複数の遺伝子ネットワーク図を保存し、それを連続的に表示するようにした請求項1記載の遺伝子ネットワーク研究支援システム。

【請求項17】 ユーザからの問合せパターンであるグラフ及び該グラフに関する付随情報を入力する手段と、前記グラフと同型なグラフを前記グラフ再構成手段によって生成されたグラフ中から検索するための検索手段と、該検索手段による検索結果に対して付随情報についての照合を行う照合手段とをさらに備えた請求項1記載の遺伝子ネットワーク研究支援システム。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】 本発明は、データベースに蓄積された遺伝子に関する情報から遺伝子ネットワークを

(3) 開2002-91991 (P2002-91991A)

構築するのを支援するための遺伝子ネットワーク研究支援システム及び方法に関する。

【0002】

【従来の技術】近年における各種生物のゲノム情報の研究作業において遺伝子の機能や発現制御などの相互関係の解明が急務となっている。そこで、この相互関係の解明のために遺伝子制御ネットワークの同定作業が行われているが、従来では、個々の遺伝子に対する機能解析実験と網羅的解析実験とは独立に行われており、これらの実験結果を統合して遺伝子ネットワーク情報を図ることができるようにしたシステムは未だ存在していない。

【0003】

【発明が解決しようとする課題】しかし、今後研究が進むにつれて網羅的解析実験が多数行われ、その結果を示すデータの蓄積が増大することが予想される。すなわち、遺伝子制御関係やタンパク間相互作用等の機能解析実験結果、転写因子予測や遺伝子機能解析等の配列情報解析結果、及び遺伝子の振る舞いや発現プロファイル実験等の網羅的解析実験結果などの、多種、多様の大量の実験データ及び膨大な情報解析結果などのデータの蓄積が予想されている。

【0004】したがって、研究者は、これらのデータを基にある刺激などに対する反応系において、それに関連する遺伝子群および最も重要な遺伝子は何か等の重要な遺伝子の同定及び着目する遺伝子を制御するネットワークを知ることで、刺激に対する反応などの生物学的現象を転写レベルで理解するため着目する遺伝子近傍のネットワークの同定等の作業を行うことが要請されるが、このような作業は膨大な時間を必要とするものである。

【0005】本発明の目的は、データベースに蓄積された遺伝子に関する多種、多様なデータに基づいて遺伝子ネットワークを構築することを支援するための、遺伝子ネットワーク研究支援システム及び方法を提供することにある。

【0006】

【課題を解決するための手段】上記課題を解決するため、本発明は、ユーザーが容易に所望の遺伝子情報を重ねて導入して遺伝子ネットワークを再構築できる遺伝子ネットワーク編集手段を設けたものである。

【0007】発明によれば、コンピュータを用い遺伝子に関する情報を統合して遺伝子ネットワークとして視覚的に表示することにより遺伝子研究を支援するシステムであって、遺伝子、RNA、タンパク質、タンパク質複合体、低分子、環境条件、細胞周期、細胞タイプ、表現型などの要素情報を要素間の結合や制御などの関係情報を蓄積したデータベースと、前記データベースに蓄積される前記要素情報をノード、前記関係情報をエッジとして表現される連結成分を生成するグラフ再構成手段と、前記グラフ再構成手段によって生成された連結成分をグラフィカルに表示する表示手段と、前記表示手段に

よって表示されているグラフのトポロジーを変更するグラフ編集手段とを備えたことを特徴とする遺伝子ネットワーク研究支援システムが提案される。

【0008】この構成によれば、例えば、細胞内あるいは細胞間の生体物質間の相互作用に関する要素情報及び関係情報を蓄積したデータベースを構築しておき、ユーザーの要求に応じて動的にそのデータベースを参照して遺伝子ネットワークを再構成して、ユーザーがそれをグラフィカルな画像として確認したり、インターフェイスに編集できるようにすることによって、遺伝子ネットワークに関する研究を支援することができる。

【0009】ここで、前記グラフ再構成手段によって生成された複数の連結成分から特定の連結成分だけを選択するユーザー入力を受け取る手段をさらに備え、これにより、特定の連結成分だけを選択的に表示できるように構成することもできる。

【0010】また、グラフ再構成手段に対して、連結成分を生成する過程で関係同士を結び付けるか否かの条件・制約を指定することができるように構成することもできる。

【0011】さらに、演繹推論により、要素間の関係に介在する要素、関係を取り除く要素間関係推論を行う演繹推論部を含むように構成することもできる。ここで、前記要素間関係推論の結果が表示されている場合に、要素間関係に対応するエッジを指定することにより、指定されたエッジに対応する要素間関係に介在する関係情報を対応するエッジを表示するようにすることもできる。

【0012】また、既存の文献あるいはデータベースから情報を抽出する情報抽出手段と該情報抽出手段によって得られた情報を前記データベースに蓄積する情報保存手段とをさらに含むように構成することもできる。

【0013】また、複数のグラフを重ね合わせてレイヤー表示するレイヤー表示手段を備えるように構成することもできる。

【0014】ここで、前記レイヤー表示手段によって表示される複数のレイヤー間で重複する又は重複しない点・辺を強調表示するようにすることもできる。

【0015】また、ユーザーからの問合せパターンであるグラフ及び該グラフに関する付随情報とを入力する手段と、前記グラフと同型なグラフを前記グラフ再構成手段によって生成されたグラフの中から検索するための検索手段と、該検索手段による検索結果に対して付随情報についての照合を行う照合手段とをさらに設ける構成とすることもできる。

【0016】前記データベースに蓄積された要素情報および関係情報に付随する情報に応じて、それらに対応する点、辺の色の濃度を変更するようにすることもできる。

【0017】付随する情報に応じて、点、辺に対応する表示プリミティブの視覚的特徴を変更する手段をさらに

(4) 開2002-91991 (P2002-91991A)

備えるようにすることもできる。

【0018】遺伝子、タンパク質が出現する文献情報を取得するステップと、関係を構成する遺伝子、タンパク質間で重複する文献情報を求めるステップと、得られた結果を表示するステップを備えるようにすることもできる。

【0019】前記データベースに各要素に付随する情報を予め付随遺伝子情報として蓄積しておくと共に、点あるいは辺を指定するステップを設けておき、指定された点あるいは辺に対応する要素の付随遺伝子情報を前記データベースから取得し、取得された付随遺伝子情報を表示するようにすることもできる。

【0020】関係情報を追加、削除するステップを設けておき、追加、削除された関係情報を加味してグラフを再構成し、再構成されたグラフを表示するようにすることもできる。

【0021】本発明によれば、また、複数の遺伝子ネットワーク図を保存し、それを連続的に表示する機構を備えた遺伝子ネットワーク研究支援システムが提案される。

【0022】

【発明の実施の形態】以下、図面を参照して本発明の実施の形態の一例につき詳細に説明する。

【0023】図1は、本発明による遺伝子ネットワーク研究支援システムの実施の形態の一例を示すシステム構成図である。遺伝子ネットワーク研究支援システム1は、外部記憶装置A、入力装置B、中央演算処理装置C及び表示装置Dを備えて成るコンピュータを用いて構成されているもので、細胞内あるいは細胞間の生体物質間の相互作用に関する情報を蓄積したデータベースを外部記憶装置Aに構築しておき、中央演算処理装置Cにおいてユーザの要求に応じて動的にそのデータベースを参照して遺伝子ネットワークを再構成して、ユーザがそれを表示装置Dによってグラフィカルな画像として確認したり、入力装置Bと表示装置Dとを用いてインタラクティブに編集できることによって、遺伝子ネットワークに関する研究を支援することができる構成となっている。

【0024】遺伝子ネットワーク研究支援システム1は、コンピュータの外部記憶装置A内に、遺伝子、タンパク質、タンパク質複合体、低分子などの細胞を構成する基本的要素に関する要素情報と、遺伝子間の制御関係、遺伝子間の類似関係、タンパク質間の相互作用などの要素間の関係を表す関係情報と、各要素に付随する情報である付随遺伝子情報とが予め蓄積されている遺伝子情報データベース(遺伝子情報DB)2が設けられている。なお、温度などの環境からの刺激あるいは細胞周期などの条件等を要素情報として加えることができる。

【0025】遺伝子情報データベース2には、遺伝子やタンパク質などの要素の名前や機能、細胞内における局

在、生物種、シノニム、発現情報、変異体などの情報が蓄積されている。また、関係情報についても、その関係の意味、関係が得られた実験とその条件などの有用な情報がデータベースに蓄積されている。これらの情報は、要素や関係の意味を理解する上で不可欠な情報であり、付随遺伝子情報として遺伝子情報データベース2内に格納されている。

【0026】遺伝子ネットワーク研究支援システム1は、遺伝子情報データベース2内に蓄積されている情報を繋ぎ合わせることにより要素間の関係をグラフとして再構成することができるようにするため、必要な情報が遺伝子情報取得手段3によって遺伝子情報データベース2から取り出され、遺伝子情報取得手段3によって取り出された情報はグラフ再構成手段4に送られる構成となっている。

【0027】遺伝子情報取得手段3及びグラフ再構成手段4によって実施されるグラフ再構成について図2及び図3を参照して説明する。図2には遺伝子情報データベース2内に蓄積されている要素情報と関係情報との一例が示されており、ここでは、要素であるタンパク質AからGまでの7つの要素情報と、これらの要素間の関係を示す関係情報とが格納されている例が示されている。要素間の関係としては、結合、活性化、不活性化の3つを例示しているが、これらに限定する趣旨ではない。

【0028】遺伝子情報取得手段3によって遺伝子情報データベース2から取り出された図2に示される情報に基づき、連結成分を生成するためのグラフ再構成手段4では図3に示されるように、連結成分を生成する。本実施の形態では、2つの連結成分が生成されている。このように、グラフ再構成手段4は遺伝子情報取得手段3によって取り出された情報に基づいて連結成分を生成し、グラフ再構成手段4はその生成結果を示すグラフ再構成データDAを再構成グラフメモリ5に出力し、再構成グラフメモリ5にこのグラフ再構成データDAが格納される。

【0029】再構成グラフメモリ5に格納されるグラフ再構成データDAの例が図5に示されている。図5に示す例では、一方の連結成分の連結成分IDが1で、他方が2となっている。そして、要素情報に関しては、7つの要素に対して要素ID1～7が付され、それぞれの要素について、ラベル、機能、細胞内局在、生物種、シノニムの各情報が付されている。一方、関係情報に関しては、連結成分IDが1の場合について関係IDが00001～00004まで設定されており、連結成分IDが2の場合については関係IDが00005～00006まで設定されている。そして、関係IDで特定される6つの関係のそれについて、始点、終点を要素IDとラベルとで特定し、その関係タイプが示されている。さらに、各関係について信頼度、情報の由来を示すデータが付加されている。ここで、要素情報、機能、細胞内局

(5) 開2002-91991 (P2002-91991A)

在、生物種、シノニムは各要素に対する付随遺伝子情報である。一方、関係情報中、信頼度及び情報の由来は各関係に対する付随遺伝子情報である。

【0030】このように、グラフ再構成手段4において、グラフ再構成データDAには各連結成分に対して連結成分IDが識別子として付与されているので、再構成グラフメモリ5において複数個の連結成分のデータの保存が可能である。グラフ再構成データDAは後述するように処理され、遺伝子ネットワークとしてディスプレイ装置6によってグラフィカルに表示されることになる。図5に示されるような形態で再構成グラフメモリ5に保存されているグラフ再構成データDAは、人間にとって理解しやすいように、図4に示すような形態でグラフィカル表示される。

【0031】図4に示されているグラフィカル表示では、各要素を点（ノード）に対応させ、各要素間の関係を辺（エッジ又はアーケ）に対応させているが、ディスプレイ装置6においてこれらの点及び辺をどのように配置するのかは一通りには決まらない。遺伝子ネットワーク研究支援システム1では、人間にとって理解しやすいよう点・辺の配置を所定の基準に基づいて決定するため、点と辺の自動配置処理を行うグラフ配置手段7が設けられており、再構成グラフメモリ5内に格納されているグラフ再構成データDAはグラフ配置手段7に送られる。

【0032】この種の自動配置の手法としては、従来から様々なものが既に提案されている。人間にとって理解しやすい、あるいは見やすいネットワークのレイアウト（配置）は、各ユーザーの好みや目的によって異なる。本実施の形態では、これら既存の複数個の手法をグラフ配置手段7内に予め備えておくこととし、ユーザーが例えば入力デバイス11を用いてこれらの手法の中から適宜選択して使用することで、ユーザーにとって理解しやすいネットワークレイアウトを得ることができるように構成されている。

【0033】グラフ配置手段7において決定されたネットワークの配置状態を示すグラフ配置データDBは配置グラフメモリ8内に一括格納される。配置グラフメモリ8に格納されたグラフ配置データDBは、図6に示されているように、再構成グラフメモリ5に格納されているグラフ再構成データDAの各要素（すなわち点）を2次元上に配置する場合の位置情報と、それらの間の関係に対応する辺を配置する場合の線の位置情報（始点と終点の位置、あるいは始点とその向きと長さ）である。辺の両端の位置は点の位置に対応するので、辺の位置情報を点の位置情報とは別々に記憶しなくても構わない。このようにする理由は、その点の位置と辺の位置が意図的に異なる位置に配置される場合に対応するためである。配置グラフメモリ8に格納されるグラフ配置データDBはグラフ表示手段9を介して表示メモリ10に出力され、

ディスプレイ装置6でグラフィカル表示される。このときの表示形態は図4に示した如くなる。

【0034】本発明による遺伝子ネットワーク研究支援システム1は、このようにしてディスプレイ装置6に表示された遺伝子ネットワーク図のトポロジーを変更して編集処理を施す機能を有している。ここで、編集とは、表示された遺伝子ネットワーク図に対してユーザーが行う点・辺の移動、削除および追加を意味すると共に、このときそれらに付随する付随遺伝子情報の編集を行うことを意味する。

【0035】次に、上述した編集機能を実現するための構成について説明する。符号11で示されるのはユーザーからの編集コマンドの入力に使用される入力デバイスで、マウス、キーボード、ポインティングデバイス等から成っている。ユーザーがマウス、キーボード等を操作することによって入力デバイス11から入力された編集コマンドは、入力デバイス処理手段12によって検知され、検知された編集コマンドは入力コマンド処理手段13によって編集コマンドとして解釈され、グラフ編集手段14に送られる。

【0036】グラフ編集手段14は、配置グラフメモリ8に格納されているレイアウト処理されたグラフ配置データDBを参照し、現在ディスプレイ装置6に表示されている遺伝子ネットワーク図を構成する点・辺の配置情報を取得する。そして、点・辺の配置情報と編集コマンドとからネットワークの内容の変更に関する第1更新データDA1を作成し、このようにして作成された第1更新データDA1は再構成グラフメモリ5に送られる。一方、ネットワークの配置変更に関する第2更新データDA2は配置グラフメモリ8に送られる。

【0037】第1及び第2更新データDA1、DA2により再構成グラフメモリ5及び配置グラフメモリ8の内容が変更され、この結果、グラフ編集手段14で編集された遺伝子ネットワークがディスプレイ装置6においてグラフィカル表示されることになる。

【0038】ここで、編集処理は、グラフのレイアウト（点・辺の位置や向き）を変更する処理をも含んでいる。この場合には、グラフ編集手段14における編集結果は配置グラフメモリ8に出力され、再構成グラフメモリ5の内容には変更が加えられず、配置グラフメモリ8の出力に従う配置状態の遺伝子ネットワーク図がディスプレイ装置6において表示されることになることは前述の通りである。

【0039】次に、図7を参照して、グラフ編集手段14の構成をより詳細に説明する。

【0040】本実施の形態においては、グラフ編集手段14は編集のための以下の5つの機能を達成することができるよう構成されている。（1）点あるいは辺の選択。（2）選択された点あるいは辺の位置や向きの変更。（3）選択された点あるいは辺の削除。（4）点あ

(6) 開2002-91991 (P2002-91991A)

るいは辺の新たな追加。（5）選択された点あるいは辺の付随情報の編集（追加、削除、変更）。

【0041】このため、上記に対応した5種類の編集用コマンドが用意され、ユーザは入力デバイス11から所要のコマンドを入力することができ、入力されたコマンドは入力デバイス処理手段12、入力コマンド処理手段13を介してグラフ編集手段14に入力される。

【0042】これら5種類の編集用コマンドは以下の通りである。

1. 編集対象となる点あるいは辺を選択するための点・辺選択コマンド（第1コマンド）
2. 選択された点あるいは辺の位置や向きを変更するための点・辺移動コマンド（第2コマンド）
3. 選択された点あるいは辺を削除するための点・辺削除コマンド（第3コマンド）
4. 点あるいは辺を新たに追加するための点・辺追加コマンド（第4コマンド）
5. 選択された点あるいは辺の付随情報を編集（追加、削除、変更）するための付随遺伝子情報編集コマンド（第5コマンド）

【0043】入力コマンド処理手段13を介してグラフ編集手段14に送られた編集用コマンドは、コマンド処理選択手段500に入力され、ここで入力された編集用コマンドの種別が識別され、そのとき入力された編集用コマンドは、この識別の結果に基づいて対応したコマンド処理部に送られる。すなわち、第1コマンドは点・辺選択コマンド処理部511に送られ、第2コマンドは点・辺移動コマンド処理部512に送られ、第3コマンドは点・辺削除コマンド処理部513に送られ、第4コマンドは点・辺追加コマンド処理部514に送られ、第5コマンドは付随遺伝子情報編集コマンド処理部515に送られる。ここで、第2、第3及び第5コマンドによる編集の処理のためには、予め第1コマンドによって編集の対象となる点あるいは辺が選択されていなければならず、したがって、第2、第3及び第5コマンドに対しては、常に第1コマンドの入力が必要となる。

【0044】先ず、第1コマンドの処理について説明すると、入力デバイス処理手段12は、ディスプレイ装置6上の任意の位置を指定するためのデバイスである2次元ポイントティングデバイス（図示せず）からの入力を検知し、その信号を2次元位置算出部501に常時出している。2次元位置算出部501は2次元ポイントティングデバイスが指定する2次元位置指定データDPを2次元位置メモリ502に出力し、2次元位置指定データDPは2次元位置メモリ502に一括格納される。点・辺選択コマンド処理部511は、第1コマンドがコマンド処理選択手段500から入力されたことに応答して2次元位置メモリ502から2次元位置指定データDPを読み取る。一方、点・辺配置情報取得手段503は、配置グラフメモリ8に格納されているグラフ配置データDB

からディスプレイ装置6に表示されている遺伝子制御ネットワークの構成要素の位置情報を読み取り、点・辺選択コマンド処理部511では点・辺配置情報取得手段503による読み取り情報と2次元位置メモリ502からの2次元位置指定データDPとからポイントティングデバイスで指定される点・辺を検索し、その結果を被選択点・辺メモリ504に出力する。

【0045】このようにして、ディスプレイ装置6上にグラフィカル表示されている遺伝子制御ネットワークを構成する所望の点又は辺をポイントティングデバイスで選択し、このようにして選択された点又は辺を特定するための被選択データDQが第1コマンドの入力に応答して、被選択点・辺メモリ504に格納される。

【0046】次に、上述の如くして選択された点又は辺を第2コマンドに従って移動させるための編集処理について説明する。点・辺移動コマンド処理部512は、点・辺の移動指令情報を含む第2コマンドをコマンド処理選択手段500から受け取る一方、現在選択されている点あるいは辺を特定するための被選択データDQを被選択点・辺メモリ504から読み取り、現在選択されている点あるいは辺を第2コマンドの内容に従って移動するよう点・辺配置情報を変更するための処理を行う。この処理結果は点・辺配置情報更新手段516に出力され、点・辺配置情報更新手段516は配置グラフメモリ8内に格納されている点あるいは辺の位置・向きの情報を変更するための第2更新データDA2を出力する。配置グラフメモリ8においては、第2更新データDA2に応答して配置グラフデータの点・辺配置情報が変更される。

【0047】この結果、ディスプレイ装置6上にグラフィカル表示されている遺伝子制御ネットワークの所望の点又は辺の位置や向きを変更する移動処理が実行される。この編集機能を用いることにより、ディスプレイ装置6上に表示されているネットワーク図を着目した要素や関係を中心とするユーザにとって見やすい表示状態に変更することができるので、ネットワークの解析等の作業効率を著しく向上させるのに役立つ。

【0048】次に、第1コマンドに従って選択された点又は辺を第3コマンドに従って削除するための編集処理について説明する。点・辺削除コマンド処理部513は第3コマンドの入力に応答し、現在選択されている点あるいは辺の情報を示す被選択データDQを被選択点・辺メモリ504から読み取り、その削除処理を行う。点・辺削除コマンド処理部513からの処理結果は再構成グラフ更新手段517に送られ、ここで、ポイントティングデバイスにより指定された点・辺のデータの削除のための第1更新データDA1が出力される。再構成グラフメモリ5においては、第1更新データDA1に応答し、被選択データDQによって示される辺または点のデータが削除される。この結果、再構成グラフメモリ5からの出力に基づいてグラフ配置手段7によって再びグラフ配置

(7) 開2002-91991 (P.2002-91991A)

処理が実行されることにより、指定された点又は辺が表示されないことになる。この編集機能を使うことにより、表示されたネットワーク内においてユーザの研究に重要でない又は必要でない要素や関係を簡単に削除することができるので、必要なものだけを見やすく表示させることができる。

【0049】点・辺追加コマンド処理部514は、追加するものが点か辺かの情報をコマンド処理選択手段500から受け取った第4コマンドの内容に基づいて判断し、点又は辺の追加を処理する。この追加処理もまた、点・辺追加コマンド処理部514での処理結果を再構成グラフ更新手段517に送り、再構成グラフメモリ5内のデータに追加すべき点又は辺のデータを加えるという更新処理によって達成される。この処理の結果、追加された点又は辺がディスプレイ装置6上に表示されるようになる。この編集機能を使うことにより、例えば遺伝子情報データベース2内には蓄積されていない自分の研究成果による情報を附加して遺伝子制御ネットワークを再構成することができるなど、外部情報を簡単に取り入れることができるようになる。

【0050】最後に、第1コマンドによって選択された点又は辺について付随する情報の追加、削除、変更等の編集のための第5コマンドに対する処理について説明する。付随遺伝子情報編集コマンド処理部515は第5コマンドの入力に応答し、現在選択されている点あるいは辺を示す被選択データDQを選択点・辺メモリ504から読み取り、その点あるいは辺に付随する情報の編集（追加、削除、変更）を処理する。付随遺伝子情報の編集処理は、被選択データDQによって示される点又は辺について、第5コマンドによって示される付随遺伝子情報を表示させるようにするための処理である。図5に示したように、再構成グラフメモリ5に格納されているグラフ再構成データDAには付随遺伝子情報が含まれており、その内容は要素ID及び関係IDを用いて特定できる。したがって、本実施の形態では、付随遺伝子情報編集コマンド515から再構成グラフ更新手段517にどの要素又は関係について付随遺伝子情報を表示させるかの情報が送られ、再構成グラフ更新手段517はこれに従って再構成グラフメモリ5内のデータのうち所要の付随遺伝子情報がグラフ配置手段7に送られるようにするための処理を実行する。この機能によれば、グラフィカル表示したネットワーク図中に要素及び又は関係について適宜に必要な文字情報を附加したりすることができる。ネットワークが説明的となり、理解しやすいものとすることができます。

【0051】図8には、本発明の他の実施の形態が示されている。図8に示されている遺伝子ネットワーク研究支援システム1Aの基本構成は図1に示した遺伝子ネットワーク研究支援システム1の基本構成と同一である。したがって、図8の各部のうち図1の各部に対応する部

分には同一の符号を付してそれらの説明を省略する。

【0052】遺伝子ネットワーク研究支援システム1Aは、下記の（1）～（8）の点で遺伝子ネットワーク研究支援システム1と異なっている。

【0053】（1）遺伝子情報データベース2から取得された複数の連結成分のリストを表示し、ユーザがそのリストから表示・編集したい連結成分を選択できるようするため、連結成分選択部20が設けられている。連結成分選択部20は、図9に詳細に示されるように、再構成グラフ選択指定手段21と、選択再構成グラフメモリ22とから成っている。

【0054】遺伝子ネットワーク研究支援システム1Aにおいては、グラフ再構成手段4において遺伝子情報データベース2内に格納されている複数個の関係情報から連結成分を生成する処理が実行されており、生成された連結成分はグラフ再構成データDAとして再構成グラフメモリ5を介して再構成グラフ選択指定手段21に送られ、例えば図10に示されるようにリスト表示される。この結果、ユーザは入力デバイス11から表示、編集したい連結成分を選択できる。

【0055】マウス、キーボード等の入力デバイス11から所望の連結成分を選択するための選択コマンドが入力されると、この選択コマンドは入力デバイス処理手段12、入力コマンド処理手段13を介して再構成グラフ選択指定手段21に入力され、所望の連結成分が選択され、選択された連結成分のみが選択再構成グラフメモリ22に送られる（図9参照）。

【0056】この選択された連結成分のデータがグラフ再構成データDAとしてグラフ配置手段7に送られ、以後、図1に示した構成の場合と同様にして、この選択された連結成分に従う選択再構成グラフが得られ、これがディスプレイ装置6においてグラフィック表示される。連結成分選択部20を設けることにより、上述の如く、遺伝子情報データベース2に基づいて生成された複数の連結成分のリストを表示し、ユーザがそのリストから表示・編集したい連結成分のみを選択することができる。図11には、上述した連結成分選択の構成の概念が示されている。

【0057】（2）連結成分を生成する過程で、関係同志を結び付けるかどうかを判断するための条件・制約をユーザが指定できるようにすることで、生成される連結成分の個数の絞りこみを行えるようにするため、条件・制約指定部30が設けられている。

【0058】連結成分選択の構成では、複数個の連結成分が選択、生成される構成であったが、遺伝子情報はその数が膨大であり、関係情報の個数はさらに膨大になることが予想される。したがって、（1）で説明した連結成分生成の結果も膨大な数になる可能性が大きい。ところが、ユーザが着目しようとする連結成分はそれらの一部にすぎず、膨大な個数の連結成分の中から、所望の連

(8) 開2002-91991 (P2002-91991A)

結成分だけをユーザが選択することは实际上困難な場合が多いと考えられる。

【0059】条件・制約指定部30はこの問題を解決することができるもので、図12に示すように、入力コマンド処理手段13とグラフ再構成手段4との間に再構成条件指定手段31と再構成条件メモリ32とを図示の如く設けた構成となっている。再構成条件指定手段31は、ユーザによって入力デバイス11から入力されるグラフ再構成の条件・制約に関するデータを受け取ってコンピュータ内で処理できる形式に変換する処理を行い、変換処理されたグラフ再構成の条件・制約に関する指定データDCを再構成条件メモリ32に出力するものである。再構成条件メモリ32に一担格納された指定データDCはグラフ再構成手段4に入力される。グラフ再構成手段4では再構成条件メモリ32からの指定データDCを考慮して連結成分のためのグラフ再構成処理が実行される。条件・制約の例としては、「グラフを再構成するために始点と終点となる遺伝子の指定」、さらにそれに加えて「再構成処理において探索する深さのレベルの制限」、あるいは「つなぎ合わせる2つの関係が満たすべき条件の指定」など様々なものと考えられる。上述の如く条件・制約指定部30を設けることにより、連結成分を生成する過程で、関係同士を結び付けるかどうかを判断するための条件・制約をユーザが指定できるようになると、生成される連結成分の個数が多い場合、ユーザに大きな負担を掛けることなしに連結成分の絞りこみが行えるようになる。

【0060】(3) 演繹推論を適用して遺伝子ネットワークを単純化すると共に、この演繹推論によって直接関係のない要素間の関係を生成することを可能にするため、演繹推論部40が設けられている。

【0061】遺伝子情報データベース2に蓄積されている情報の量は厖大である。また、グラフ再構成手段4によって再構成される遺伝子ネットワークもまた、厖大な要素情報と関係情報とから構成される複雑なものになる可能性がある。今後、遺伝子情報データベース2に蓄積される情報は増加するため、この傾向は強くなるはずである。ネットワークが複雑になると、たとえグラフィカルに表示したとしても、人間が理解することは困難である。

【0062】演繹推論は、既知の知識の組から関連する別の表現を導き出すものである。図13を使って具体的に説明する。各記号の意味は凡例に示した通り、四角が遺伝子、円形がタンパク質、辺(矢印等)が関係を表している。今、図13のような遺伝子情報があるとする。図13では、分かりやすさのために遺伝子情報を模式的に表現している。それらの関係を言葉で表すと、「遺伝子Aがタンパク質Aをコードする」、「遺伝子Bがタンパク質Bをコードする」、「遺伝子Cがタンパク質Cをコードする」、「タンパク質Bがタンパク質Aを活性化

する」、「タンパク質Cがタンパク質Aを不活性化する」、「タンパク質Aが遺伝子Cの転写を促進する」となる。いま、推論規則として次の2つ、すなわち(1)「タンパク質1がタンパク質2を活性化し、かつタンパク質2が遺伝子3の転写を促進するならば、タンパク質1が遺伝子3の転写を促進する」、(2)「遺伝子1がタンパク質1をコードし、かつタンパク質1が遺伝子2の転写を促進するならば、遺伝子1が遺伝子2の転写を促進する」を適用することにより、「遺伝子Bが遺伝子Cの転写を促進する」なる関係が新たに得られる。この演繹推論により得られた関係は、介在している他の遺伝子やタンパク質を含んでいない。そのため、記述が簡潔であり、理解が容易である。

【0063】また、図14に示したように、グラフ再構成により得られたグラフに対してタンパク質間の類似関係を考慮することによって、例えば新たなタンパク質間の活性化の関係を演繹推論によって得ることが可能である。

【0064】演繹推論部40は上述の如く演繹推論を適用して遺伝子ネットワークを単純化すると共に、この演繹推論によって直接関係のない要素間の関係を生成することを可能にするもので、図15に詳細に示されるように、上記演繹推論のための推論規則が蓄積されている推論規則データベース(推論規則DB)41と、入力デバイス11からユーザが入力した選択コマンドが入力コマンド処理手段13を介して入力されている推論規則選択手段42とを備えている。推論規則選択手段42は選択コマンドに従って推論規則データベース41からユーザによって選択された推論規則データDDを取り出して推論規則メモリ43に一担格納し、その後推論規則データDDは推論規則メモリ43から推論手段44に送られる。

【0065】推論手段44では、再構成グラフメモリ5に格納されているグラフ再構成データDAに対して推論規則メモリ43に格納されている推論規則データDDに基づく推論処理を施し、図13及び図14に基づいて説明したような所要の推論処理を施された推論結果データDEが推論結果メモリ45に格納される。推論結果メモリ45に格納された推論結果データDEはグラフ配置手段7及びグラフ編集手段14に送られる。グラフ配置手段7では、再構成グラフメモリ5のグラフ再構成データDAと推論結果メモリ45の推論結果データDEとを合わせることによりグラフの単純化等の目的とする処理を行い、グラフの配置処理を行う。以後のグラフィカル表示のための処理は、図1に示した遺伝子ネットワーク研究支援システム1の場合と同様である。

【0066】上述の如く演繹推論部40を設けると、厖大な要素情報と関係情報とから構成される複雑な遺伝子ネットワークを単純化すると共に、直接関係のない要素間の関係を生成することを可能にし、理解しやすいグラ

(9) 開2002-91991 (P2002-91991A)

フ表示をすることができる。

【0067】(4) 遺伝子ネットワーク研究支援システム1Aにおいては、上記(3)の構成による多段推論(本明細書では、複数回の推論が組み合わされるような推論をこう呼ぶこととする)によって得られた最終的なグラフの配置処理が行われたグラフだけが表示されているディスプレイ装置6をマウスでクリックすると、その推論の段数が辺の側に表示される。さらに、その推論結果の基になった関係に対応する辺が太く表示される。それと同時に、その推論の基になった推論結果に対応する関係の辺がその段数と共に表示される。そこで、その基になった推論結果に対応する辺をクリックすると、さらにその推論の基になった関係に対応する辺が太く表示される。

【0068】異なる推論を繰り返し適用することによって、遺伝子ネットワークをより単純化することができるが、このように推論結果によって求められた関係が、実は他の推論結果に基づいて求められている場合が起こり得る。例えば、図16には2つの推論を適用することによって、遺伝子間の制御関係を新たに推論する例が示されている。図16では、「タンパク質Aがタンパク質Cを活性化する」、「タンパク質Eがタンパク質Dを活性化する」、「タンパク質Eが遺伝子Bの転写を促進する」という関係があるが、遺伝子Aと遺伝子Bの間に制御関係を推論することはできない。そこで、タンパク質間の類似関係に着目し、類似しているタンパク質同士の機能が類似するという仮説を設け、それに基づいて推論を行う。すなわち、「タンパク質Cとタンパク質Eが類似する」という関係に着目し、「タンパク質Eがタンパク質Dを活性化すること」から、「タンパク質Cがタンパク質Dを活性化する」という関係を推論する。次に、新たに推論された「タンパク質Cがタンパク質Dを活性化する」という関係から、結局「遺伝子Aが遺伝子Bの転写を促進する」という関係が得られる。

【0069】このような場合、得られた推論結果がどのような推論のプロセスに基づいた結果であるのかが分からぬ。また、推論規則として信憑性の乏しいものが使われていることも考えられる。この場合は、推論結果が予測的であるため、推論結果を得るプロセスを知ることが不可欠である。例えば、上記のタンパク質の類似性に基づいた推論によって得られる新たな関係は、予測的な関係である。したがって、得られた推論結果が、他の推論結果に基づいているかどうかを把握できるような遺伝子ネットワーク表示方法が必要である。

【0070】この問題を解決するようにするために、図15に示した演繹推論部40の構成においては、推論結果メモリ45に格納されている推論結果データDEをグラフ編集手段14に送り、グラフ編集手段14においてこの推論結果データDEを参照してグラフ編集を行うように構成されている。この場合、推論規則選択手段42

では、複数の推論規則を選択し、その適用順序を決定し、その決定結果である推論規則データDDを推論規則メモリ43を介して推論手段44に出力する。

【0071】推論手段44は、推論規則メモリ43から推論規則データDDを受け取ると、再構成グラフメモリ5からのグラフ再構成データDAに対してこの推論規則データDDを適用する。推論結果メモリ45内に格納されている推論結果データDEは、他のどの推論結果に基づいているのかを表す情報と、その推論結果の段数の情報を持つ。これによって、全ての推論結果の段数と他のどの推論結果に基づいているかが分かる。

【0072】次に、表示方法について説明する。推論結果メモリ45内に格納されている推論結果データDEは、図1に基づいて説明したように、配置グラフメモリ8、グラフ表示手段9を介してディスプレイ装置6に表示される。次に、ユーザが、入力デバイス11を使って、推論結果であるエッジを選択する。すると、その選択コマンドが、入力コマンド処理手段13によって解釈され、グラフ編集手段14に出力される。グラフ編集手段14は、配置グラフメモリ8を参照して、何が選択されたかを検知する。また、選択された辺が推論に基づいているかどうかを推論結果メモリ45に格納されている推論結果データDEを参照して判定する。もし、その辺が推論に基づいているのであれば、推論の段数とその基になった他の関係に対応する辺の表示形態を変更し、その結果を配置グラフメモリ8に反映させる。図17には多段推論の結果の表示例が示されている。

【0073】演繹推論部40を以上のように構成したので、得られた推論結果がどのような推論のプロセスに基づいた結果であるかを表示することができる。

【0074】(5) 再構成したグラフの構成要素に付随する情報に基づいて、点・辺の濃度を変更できるようにすることで、ユーザの信頼性の判定基準を反映した遺伝子ネットワークの表示を行えるようにするために、表示形態変更部50が設けられている。

【0075】表示形態変更部50は、図19に示されているように、図8に示した入力コマンド処理手段13とグラフ表示手段9との間に、表示形態条件指定手段51と表示形態条件メモリ52とを設けた構成によって実現することができる。表示形態条件指定手段51は、入力デバイス11からユーザが入力する表示形態条件指定コマンドを入力コマンド処理手段13を介して受け取り、表示形態条件指定コマンドに従って属性値とそれに対応する色の濃度、あるいは全体のコントラストを設定する。このような方式に基づく遺伝子ネットワーク表示例を図18に示す。前方にあるウィンドウで、属性値の対応を選択し、その属性値に対応する濃度を指定する。また、ネットワーク構成要素の信頼度が低いために、全体的に濃度が薄くなってしまう場合に備えて、全体のコントラストを変更できるようにした。表示形態条件指定手

(10) 2002-91991 (P2002-91991A)

段51ではこのようなグラフ設定データDFが設定され、グラフ設定データDFは表示形態条件メモリ52に送られ一括格納された後、グラフ表示手段9に送られる。そして、グラフ設定データDFに従ってネットワーク図が表示されることになる。この方法によれば、一度再構成したグラフの各構成要素の表示濃度を変更するだけであるので、ユーザの判断基準の変更を、高速に反映させることができる。

【0076】遺伝子情報データベースに蓄積される情報は公開データベース、文献、実験結果やバイオインフォマティクスを駆使した情報技術に基づく予測解析結果など多様な情報源から取得されることが考えられる。それらの情報の中には、信頼できる情報やそうでない情報が含まれている。また、信頼するかどうかの判定基準には必ずしも絶対的ではなく、ユーザや使用目的により多様である。したがって、一律にそれらの情報からグラフを再構成して遺伝子ネットワークとして表示することは、信頼できる情報とそうでない情報のユーザによる区別を難しくしてしまうことになる。ユーザのデータに対する信頼の度合いを反映したグラフの再構成や表示手段が必要である。

【0077】このような問題に対して、連結成分の絞り込みにより、再構成条件を指定して、一定の信頼できる情報だけを選別してグラフを再構成することが考えられる。しかしながら、それでは、グラフ再構成に使用するデータを絞り込んでしまうことになり、結果的に重要な関係を見逃してしまうことになりかねない。その場合は、絞込みを緩めて、グラフの再構成を再試行しなければならない。この方式は、グラフの再構成には比較的処理に時間を要するため、ユーザにとって効率的ではない。

【0078】最良の解決策は、ユーザの情報の判定基準を考慮したグラフ再構成と、その結果を表示する一連の処理を高速に処理できるようにして、ユーザの判断基準の変更に対して即時に結果を表示できるようにすることである。しかしながら、先述したように、遺伝子ネットワークのように膨大なデータに基づくグラフ再構成処理の高速化は困難である。

【0079】表示形態変更部50を設けることにより、再構成したグラフの構成要素に付随する情報に基づいて、点・辺の濃度を変更できるので、ユーザの信頼性の判定基準を反映した遺伝子ネットワークの表示が行えるようになる。

【0080】(6) 多様な条件で構築された遺伝子ネットワーク図を蓄積しておき、それらを要素間の類似性や関連性などを基準にして重ね合わせて表示することができるようにするためレイヤー表示部60が設けられている。

【0081】本機構に類似した既存の報告事例として、Peter Uetz, Loic Giot, Gerand Cangney, Traci A. Ma

nsfield, Richard S. Jundson, James R. Knight, Daniel Lockshon, Vaibhav Narayan, Maithreyan Srinivasan, Pascale Pochart, Alia Qureshi-Emili, Ying Li, Brian Godwin, Diana Conover, Theodore Kalbfleisch, Govindan Vijayadamodar, Meijia Yang, Mark Johnston, Stanley Fields, Jonatha M. Rothberg, "A comprehensive analysis of protein-protein interactions in *Saccharomyces cerevisiae*", NATURE, Vol.403, 10 Feb 2000がある。この中で紹介された機能は、タンパク質間相互作用関係についてタンパク質の生物種間相同性に基づいて重ね合わせて表示することができるものである。本発明においては、環境条件やグラフ再構成に適用した条件・制約の違いによる遺伝子ネットワークの変化を理解しやすく提示するための表示方式に関するものであり、図20は実現するレイヤー表示の一例であり、L101とL102とはそれぞれ異なるレイヤーである。図20では、レイヤーが3次元的に表示されているが、これを垂直方向に正射投影することによって、2次元の図として表示することも可能である。レイヤーL101とL102は、それぞれS101～S107とS201～S209の遺伝子およびタンパク質から構成される遺伝子ネットワークが配置されている。この2つのレイヤーは、それぞれ異なる環境条件に限定して再構成された遺伝子ネットワークであり、両レイヤーのS101とS201、S102とS202、S104とS204は同一の遺伝子とタンパク質に対応している。これらのことから、異なる環境条件においても、類似した遺伝子ネットワークがS101(S201)を基点として存在することが分かる。更に、S101(S201)を基点とした全く異なる遺伝子ネットワーク(S101→S105→S106→S107)、(S201→S208→S209)が存在することも分かる。

【0082】図8を参照すると、レイヤー表示部60は、再構成グラフ情報と配置グラフ情報を併せ持ったグラフ情報を保存する機能を有している。入力デバイス11を用いてユーザによって指定されたグラフ情報保存読取コマンドは、保存コマンドと読取コマンドとの2種類を有しており、いずれもグラフ情報保存読取手段61に入力される。保存コマンドの場合、グラフ情報保存読取手段61では、再構成グラフメモリ5からのグラフ再構成データDAと配置グラフメモリ8からのグラフ配置データDBとを受け取り、それらを処理してグラフ情報データDGとしてグラフ情報62に送り、格納する。また、読取コマンドの場合は、グラフ情報保存読取手段61がグラフ情報62からグラフ情報データDGを読み込み、読み込んだグラフ情報データDGをグラフ再構成データDAとグラフ配置データDBとに分割して、それぞれを再構成グラフメモリ5、及び配置グラフメモリ8に出力する。

【0083】レイヤー表示コマンドには、既存のグラフ

(11) 2002-91991 (P2002-91991A)

をどのレイヤーに対応付けるかを設定するコマンド（レイヤー設定コマンド）と、どのレイヤーを表示・非表示するかを指定するコマンド（レイヤー表示コマンド）と、レイヤー間の再構成グラフの要素を指定された関係（先の例では、同一の遺伝子、タンパク質であるという関係）で対応付け、人間にとて見やすく配置するコマンド（レイヤー配置コマンド）がある。ユーザーによって与えられたレイヤー設定コマンドはレイヤー表示手段63に入力される。レイヤー表示手段63では、再構成グラフメモリ5に属するグラフ再構成データDAの所望の再構成グラフデータに対してそれが対応するレイヤーを設定する。それと同時に、その再構成グラフメモリ5に対応する配置グラフメモリ8のレイヤーと同じレイヤーに設定する。

【0084】レイヤー表示コマンドの場合、レイヤー表示手段63は表示・非表示するレイヤーを受け取り、それをグラフ表示手段9に出力する。グラフ表示手段9は、入力された表示レイヤーに属するグラフだけを表示メモリ10に出力する。

【0085】レイヤー配置コマンドの場合、レイヤー表示手段63は、レイヤー間で対応付けに使用する関係を入力として受け取り、それをグラフ配置手段7に出力する。グラフ配置手段7は、再構成グラフメモリ5からグラフ再構成データDAを受け取り、異なるレイヤーに属するグラフ間で、指定された関係に基づいて対応付けて見やすく配置し、その結果は配置グラフメモリ8に格納される。

【0086】上述した構成によれば、多様な条件に基づいて再構成された遺伝子ネットワークをレイヤー表示機構により表示することによって、それらの違いや類似性の把握を容易にすることができる。

【0087】(7) 多様な条件下における遺伝子ネットワークを比較する上で有効な方式の一つとして、上述したレイヤー表示がある。しかし、レイヤー表示される遺伝子ネットワークが複雑になってくると、それらの違いや共通点を把握することは困難である。そこで、レイヤー表示部60の上述の機能を用いることにより複数レイヤーで表示される遺伝子ネットワークの共通部分と差分とをそれぞれ強調表示することができるようになると、上述したレイヤー表示部60のレイヤー表示手段63には所要の強調表示のための処理機能が付加されている。この強調表示のための処理機能は、既にレイヤー表示が為されている状態で、レイヤー間の差分強調表示あるいは共通部分強調表示をするためのものである。ユーザーにより入力デバイス11から入力された差分強調表示コマンド、共通部分強調表示コマンドが入力コマンド処理手段13を介してレイヤー表示手段63に入力される。

【0088】共通部分強調表示コマンドの場合、レイヤー表示手段63は、再構成グラフメモリ5のグラフ再構

成データDAと配置グラフメモリ8のグラフ配置データDBとを参照して共通部分を求める。得られたレイヤー間の共通部分の情報は、再構成グラフメモリ5と配置グラフメモリ8とに分けられ、そして新たなレイヤーに対応付けられ、それぞれ再構成グラフメモリ5と配置グラフメモリ8とに出力される。

【0089】次に、グラフ表示手段9に対して、共通部分に対応付けられたレイヤーを出し、グラフ表示手段9がそのレイヤーを表示メモリ10に出力する。

【0090】差分強調表示コマンドの場合、レイヤー表示手段63は、再構成グラフメモリ5内に格納されているグラフ再構成データDAと配置グラフメモリ8内に格納されているグラフ配置データDBとを参照してレイヤー間の差分を求める。得られたレイヤー間の差分の情報は、再構成グラフデータと配置グラフデータとに分けられ、そして新たなレイヤーに対応付けられ、それぞれ再構成グラフメモリ5と配置グラフメモリ8とに出力される。

【0091】次に、グラフ表示手段9に対して、差分グラフに対応付けられたレイヤーを出し、グラフ表示手段9がそのレイヤーを表示メモリ10に出力する。

【0092】ここで図21は、レイヤー間の共通部分と差分の強調表示機能が模式的に示されている。図21では、レイヤーが3次元的に表示されているが、前項と同様に、レイヤー面を垂直な方向から正射投影することによって、2次元的に表示することも可能である。

【0093】レイヤーL101、L102はそれぞれ環境条件をA、Bに限定して得られた遺伝子ネットワークである。2つの遺伝子ネットワークの要素に着目すると、同一の要素（S101とS201、S102とS202、S103とS203、S104とS204、S106とS206、S107とS207、S108とS208）が含まれていることが分かる。これらは、2つの異なる条件下でも、共通して反応が確認されている部分である。一方、S105、S209はそれぞれのレイヤーにのみ存在している。これは、2つの条件下において、異なる反応経路が存在していることを表している。レイヤーL103はレイヤーL101とL102の差分を強調表示したレイヤーである。レイヤーL104はレイヤーL101とL102の共通部分を強調表示したレイヤーである。

【0094】レイヤー表示手段63に上述した処理機能を付加したので、多様な条件下に対応してレイヤー表示された複雑な遺伝子ネットワークを比較し、それらの違いや共通点を容易に把握することができ、非常に有効である。

【0095】(8) 遺伝子ネットワークには、類似するパターンが存在する場合がある。例えば、図22で示すように、(1)ある遺伝子Aが別の遺伝子Bおよび遺伝子Cの転写を促進する場合、(2)ある遺伝子Aが別の

(12) 2002-91991 (P2002-91991A)

遺伝子Bの転写を促進し、その遺伝子Bが遺伝子Aの転写を抑制するなどである。ここで示した例は、遺伝子間の単純なパターンであるが、これにタンパク質間の相互作用も含めたり、環境条件や特定の遺伝子やタンパク質を介在させるなどのパターンに基づいて検索する場合も考えられる。そこで指定されたグラフのパターンを部分グラフとして持つグラフを検索することができるようになるため、検索部70が設けられている。

$$\forall i, j \{v_{\alpha i}, v_{\alpha j}\} \in E_{\alpha} \Rightarrow \{f(v_{\alpha i}), f(v_{\alpha j})\} \in E$$
 となる写像 $f: V_{\alpha} \rightarrow V$ が存在 (1)

写像 f の選び方は、 $(G_{\alpha}$ の点の個数) P (G_{β} の点の個数) 通り存在するので、グラフを構成する点の個数が増加すると実用時間では計算不可能となってしまう。本発明は、この問題を解くアルゴリズムではないので、その実装方法としては、既存のアルゴリズムを使用するものとする。アルゴリズムとしては、Ullmannのアルゴリズム (J. R. Ullmann. An algorithm for subgraph isomorphism. Journal of the ACM, 23(1):31-42, January 1976) が知られている。

【0097】図23には検索部70の詳細構成が示されている。従来の部分グラフ同型問題では点や辺に付随する情報まで考慮していないが、ここでは、まず、指定されたパターンを部分グラフとして持つグラフの検索結果(グラフと写像の組)を記憶しておく、それからグラフに対して付随情報が合っているかどうかについて詳細な検証を行うことで対応する構成となっている。

【0098】検索部70は、同型グラフ検索手段71と、問い合わせパターンメモリ72と、検索結果メモリ73とから成り、同型グラフ検索コマンドには、問い合わせグラフ指定コマンドと同型グラフ検索コマンドの2つのコマンドがある。問い合わせグラフ指定コマンドは、問い合わせしたいグラフのパターンで構成される。一方、同型グラフ検索コマンドは、再構成グラフメモリ5に蓄積されているグラフ再構成データDAに対して、問い合わせパターンメモリ72に同型なサブグラフを含むグラフの探索を行う。

【0099】まず、問い合わせグラフ指定コマンドについて説明する。ユーザーによって入力デバイス11から入力された問い合わせグラフ指定コマンドは入力コマンド処理手段13を介して同型グラフ検索手段71に入力される。同型グラフ検索手段71は入力されたグラフ指定コマンドを解析し、そこに含まれている問い合わせパターンを抽出し、抽出された問い合わせパターンを示す問い合わせパターンデータDIを問い合わせパターンメモリ72に退避させる。

【0100】次に、同型グラフ検索コマンドについて説明する。ユーザーによって入力デバイス11から入力された同型グラフ検索コマンドもまた入力コマンド処理手段13を介して同型グラフ検索手段71に入力される。これに応答して同型グラフ検索手段71は、問い合わせパターンメモリ72に退避されている問い合わせパターンデータD

【0096】あるグラフ G_{α} が別のグラフ G_{β} の部分グラフに同型であるかどうかを見つける問題は、グラフ理論の分野において部分グラフ同型問題と呼ばれている。2つのグラフ $G_{\alpha} = (V_{\alpha}, E_{\alpha})$ 、 $G_{\beta} = (V_{\beta}, E_{\beta})$ について、 G_{β} の部分グラフと G_{α} が同型とは、両グラフの全ての頂点と全ての辺に対して完全な1対1対応がある場合のことである。 $V \subseteq V_{\beta}$ 、 $E \subseteq E_{\beta}$ とすると、部分グラフ同型の条件は下式(1)のようになる。

$$\forall i, j \{v_{\alpha i}, v_{\alpha j}\} \in E_{\alpha} \Rightarrow \{f(v_{\alpha i}), f(v_{\alpha j})\} \in E$$

Iを読み出す。また、再構成グラフメモリ5に保存されているグラフ再構成データDAを読み出し、先の問合せパターンに同型な部分グラフを持つグラフとその写像の組みを検索結果メモリ73に出力する。部分グラフ同型の探索が全てのグラフ再構成データDAに対して完了したら、次に、問合せパターンに含まれている要素・関係の付随情報の検証を検索結果メモリ73に保存されている全ての検索結果に対して行う。それが完了したら、検索結果をグラフ表示手段9に出力することにより、ディスプレイ装置6に検索結果が表示される。

【0101】また、上述したレイヤー表示機能を用いて、検索結果を問合せパターンとともにレイヤー表示することも可能である。さらに、検索結果が複数個存在する場合は、それら全てをやはりレイヤー表示機能を用いてレイヤー表示することが可能である。

【0102】この検索の機能は、例えば、モデル生物で観察された遺伝子ネットワークのパターンがより高等な生物にも存在するかどうかを確認する場合に有効である。

【0103】図24には、本発明の他の実施の形態が示されている。図24に示されている遺伝子ネットワーク研究支援システム1Bの基本構成は図1に示した遺伝子ネットワーク研究支援システム1の基本構成と同一である。したがって、図24の各部のうち図1の各部に対応する部分には同一の符号を付してそれらの説明を省略する。

【0104】遺伝子ネットワーク研究支援システム1Bは、下記の(9)～(11)の点で遺伝子ネットワーク研究支援システム1と異なっている。

【0105】(9) 遺伝子ネットワーク研究支援システム1Bにあっては、付随遺伝子情報表示手段81を設け、これにより、グラフィカルに表示された再構成グラフ(遺伝子ネットワーク)の点・辺のそれぞれ対応する要素・関係に付随する情報を遺伝子情報データベース2から取得し、ディスプレイ装置6に表示できるように構成されている。

【0106】図25は、模式的に、遺伝子ネットワークと、それを構成する要素・関係に付随する情報(A101、A102、A103、A104)を並べて記載したものである。

(13) 2002-91991 (P2002-91991A)

【0107】図24を参照して上述した付随情報表示のための動作について説明すると、ユーザーによって入力デバイス11から入力された付随遺伝子情報表示コマンドは、入力コマンド処理手段13によって付随遺伝子情報表示コマンドであると解釈されて、付随遺伝子情報表示手段81に出力される。付隨遺伝子情報表示コマンドは、付隨情報の表示対象である要素あるいは関係の識別子から構成される。付隨遺伝子情報表示手段81は、入力される要素あるいは関係の識別子に対応する付隨情報を遺伝子情報取得手段3から得る。得られた付隨情報はグラフ表示手段9を介して表示メモリ10に出力され、ディスプレイ装置6で表示される。

【0108】図26には、その表示形態の一例が示されている。

【0109】遺伝子ネットワーク研究支援システム1Bにあっては、付隨遺伝子情報表示手段81を設けたので、グラフィカルに表示された再構成グラフ（遺伝子ネットワーク）の点・辺のそれぞれ対応する要素・関係に付隨する情報をディスプレイ装置6に表示でき、ユーザーは必要性に応じて要素や関係の意味を理解する上で不可欠な情報を取得できる。

【0110】（10）遺伝子ネットワーク研究支援システム1Bにおいては、さらに、要素・関係に付隨する情報に応じて、それらに対応する点・辺の色、辺種、形状等の表示形態を変えることができるようするため、表示形態条件指定手段82及び表示形態条件メモリ83が設けられている。

【0111】グラフィカルに表示される再構成グラフは、点・辺で表される。それらの意味を理解するには、上述の如く、それらに対応する要素・関係に付隨する情報を参照することが有効である。ここでは、付隨する情報を応じて再構成グラフを構成する点・辺の表示形態を変更できるようにすることによって、表示された遺伝子ネットワーク図から直観的に意味を理解できるようにすることを意図している。

【0112】表示形態条件指定手段82はユーザーによって入力デバイス11から入力された表示形態変更条件を示す表示形態変更条件データDHを表示形態条件メモリ83に出力し、表示形態条件メモリ83に一括格納する。表示形態変更条件データDHは表示形態条件メモリ83からグラフ表示手段9に送られ、グラフ表示手段9においては表示形態条件メモリ83から受け取った表示形態変更条件データDHに基づいて遺伝子ネットワーク図を作成した後、この遺伝子ネットワーク図を表示メモリ10に送り、これにより遺伝子ネットワーク図がディスプレイ装置6において表示される。

【0113】図25にはその表示例が示されている。この例では、転写因子が太線の白抜き円形、タンパク質複合体が黒丸、それ以外のタンパク質が細線の白抜き円形、DNA塩基配列が模式的な画像で表されている。ま

た、タンパク質同士の結合関係が1点鎖線、タンパク質複合体とその構成要素間の関係が破線、タンパク質がDNA塩基配列に結合する関係が実線で表されている。

【0114】DNA塩基配列の例では、点に対応する図形の形状の表示形態を指定するのではなく、予め用意された画像データを表示している。すなわち、ここでは、表示アリミティの表示形態として、画像を表示することも含めている。

【0115】表示形態条件指定手段82及び表示形態条件メモリ83を設けることにより、要素・関係に付隨する情報に応じて、それらに対応する点・辺の色、辺種、形状等の表示形態を変えることができるようし、表示された遺伝子ネットワーク図から直観的に意味を理解できるようにしている。

【0116】（11）遺伝子情報データベース2を構築するには、既存の公開データベースから必要な情報を抽出してデータベースに蓄積する方法が現実的である。その既存の公開データベースには、遺伝子（あるいはそれがコードしているタンパク質）に関する様々な情報が対応付けられており、それらの情報の中には、その遺伝子（あるいはタンパク質）に関する過去の論文（一般に複数）の情報が記載されている場合が多い。遺伝子ネットワーク研究支援システム1Bにおいては、遺伝子やタンパク質等の要素間の関係情報が信頼できるかどうかの度合いをその関係を構成している要素（2項関係ならば2つの要素）が参照する文献情報に基づいてその信頼度を数値化して表示できるように構成されている。

【0117】遺伝子情報データベース2を参照して、再構築された遺伝子ネットワークには、多様な情報源に基づいた情報がモザイク状に埋め込まれている。これら遺伝子ネットワークに埋め込まれた様々な情報の中のどれを信頼するかは個々のユーザーやシステムの利用目的によって異なる。この信頼度を遺伝子ネットワークの表示に即時に反映させる方法として、既に遺伝子ネットワーク研究支援システム1Aの構成に関連して遺伝子ネットワークの構成要素に付隨する情報に基づいて表示色の濃度に反映させる方法を述べた。この方法は、遺伝子情報データベース2に登録されているような既に報告されている関係情報がどの程度信頼できるかを直観的に把握したい場合には有効である。

【0118】それとは別に、既述の如く、ユーザーが行った実験結果（例えば、タンパク質間の相互作用の抽出実験結果、遺伝子間の制御関係の抽出に関する実験）をシステムに反映させて遺伝子ネットワークとして可視化する場合、それらの関係の信頼度を遺伝子情報データベース2に蓄積されている情報を使って把握したい場合がある。既に、遺伝子情報データベース2から構築された遺伝子ネットワークと重ね合わせることにより、実験結果の検証を行うことについて説明した。しかしながら、もっとシステムティックにそれらの信頼度を数値で算出す

(14) 2002-91991 (P2002-91991A)

る方法があれば、大規模なデータの検証にとって有効である。

【0119】次に本実施の形態における関係情報の信頼度の数値化方法について説明する。関係情報が2項関係の場合について説明する。構成する要素をそれぞれA、Bとし、以下のように定義する。

n_{ab} : 要素A、Bが共通して参照している文献数
 C_i : 要素A、Bが共通して参照している文献 ($i = 1, \dots, n_{ab}$)

もし要素A、B間の関係が高いのであれば、そうでない関係よりも共通して参照している文献の個数 n_{ab} がより大きくなると考えられる。ここでのコンセプトは、要素がデータベースにおいて共通して参照する文献情報に基づいて要素間の関係の信頼度を数値化することである。ただし、要素間で共通して参照される文献には、多くの文献で参照される文献が存在する。このような文献は多くの要素が参照するため、あまり重要度が高くない。すなわち、ある要素で特異的に参照される文献が重複している場合は、それらの要素間の関係に高い関連があるとするような数値化方法が必要である。次に、文献の特異性を考慮したスコアリングの方法について説明する。

【0120】いま、

D : データベースに登録されている全要素数

d_{ci} : 文献 C_i を参照する要素の個数

とし、文献の特異性のスコア S_{ci} を、下式(2)のように定義する。

$$S_{ci} = \frac{D}{d_{ci}} \quad (2)$$

そこで、関係情報の信頼度を全ての共通して参照されている文献の特異性に対するスコア S_{ci} の和をとって下式(3)のように定義する。

$$S_c = \sum_{i=1}^{n_{ab}} S_{ci} \quad (3)$$

文献の特異性に対するスコア S_{ci} はあまり参照されない文献では極端に大きくなる場合があり、そのような場合、関係情報の信頼度のスコアの比較に強い影響を及ぼす恐れがある。そこで、この影響を抑制するために、下式(4)のように対数を取ると良い場合もある。

$$S_{ci} = \log\left(\frac{D}{d_{ci}}\right) \quad (4)$$

【0121】以上のような方法で要素がデータベースにおいて共通して参照する文献情報に基づいて要素間の関係の信頼度を数値化したので、例えばユーザが行った実験結果(例えば、タンパク質間の相互作用の抽出実験結果、遺伝子間の制御関係の抽出に関する実験)をシステムに反映させて遺伝子ネットワークとして可視化する場合、それらの関係の信頼度を遺伝子情報データベース2

に蓄積されている情報を使って把握したい場合、システムに信頼度を数値で算出することができる。

【0122】図27には本発明の第4の実施の形態が示されている。図27に示されている遺伝子ネットワーク研究支援システム1Cの基本構成は図1に示した遺伝子ネットワーク研究支援システム1のそれと同一である。したがって、図27の各部のうち図1の各部に対応する部分には同一の符号を付してそれらの説明を省略する。

【0123】遺伝子ネットワークは環境条件の変化によって変化するものであり、遺伝子ネットワークの変化を把握することは細胞内の分子機構を理解する上で極めて重要である。遺伝子ネットワーク研究支援システム1Cは変化している各遺伝子ネットワーク図を保存し、保存された遺伝子ネットワーク図を変化がわかるように連続的に表示するアニメーション機能を有している点で遺伝子ネットワーク研究支援システム1と異なっている。

【0124】上述により遺伝子情報データベースを参照して再構成される遺伝子ネットワークは、環境条件によって異なるが、このような環境条件として代表的なものに、細胞周期がある。細胞周期とは、真核生物細胞が増殖分裂する周期のことをいう。細胞周期には、G1期、S期、G2期、M期がある。G1期では、細胞はまずDNA合成の準備をする。次に、S期においてDNA合成が行われDNA量が倍加する。その後、準備期間G2期があり、それから、顕微鏡で観察されるような染色体の出現、移動、細胞分裂などが起こる。この時期がM期である。分裂を行った細胞はまたG1期に戻り、次のサイクルに入り、同じことが再度繰り返される。すなわち、G1期→S期→G2期→M期という周期で細胞が増殖するわけである。細胞によっては、増殖しないものがあるが、これはG1期からS期に行かず、G0期に入って休止してしまったものである。また、下等真核生物ではG2期からG0期に入ることもある。

【0125】次に、上述のような環境条件の変化に伴う遺伝子ネットワークの変化を把握するためのアニメーション機能を遂行するための各構成について説明する。

【0126】アニメーション機能は図27に示すように、アニメーション表示手段111とアニメーションデータメモリ112とからなるアニメーション表示処理部100を設けることにより実現される。

【0127】遺伝子ネットワーク研究支援システム1Cにおいて、アニメーション表示に関する設定コマンドには、アニメーション編集コマンドと、アニメーション表示コマンドがある。これら2つのコマンドは、入力デバイス11により入力されると入力デバイス処理手段12を介して入力コマンド処理手段13に出力され、入力コマンド処理手段13で解釈されアニメーション表示手段111に送られる。

【0128】アニメーション編集コマンドでは、表示対象となる遺伝子ネットワークの選択およびその順序を指

(15) 月 2002-91991 (P 2002-91991A)

定する。アニメーション表示コマンドは、アニメーション編集コマンドによって設定された順序で遺伝子ネットワークを順番に表示する。この結果、ユーザは入力デバイス11からアニメーション表示したい遺伝子ネットワークを選択し順序を指定することにより、設定された順序でディスプレイ装置6にアニメーション表示することができる。

【0129】アニメーション編集コマンドについて説明する。アニメーション編集コマンドを処理するために、あらかじめ表示対象となる全ての遺伝子ネットワークが再構成グラフデータとしてメモリ上に存在していかなければならない。そのために、グラフ再構成手段4で遺伝子ネットワークを複数個生成するか、あるいは、前述したようにグラフ情報保存読取手段によって、予め保存されているグラフ情報を読み込んでおく必要がある。

【0130】アニメーション編集コマンドは2つの要素から構成される。1つは、対象とする遺伝子ネットワークを指す識別子であり、もう一つは、その遺伝子ネットワークを表示する順序を示す番号である。対象とする遺伝子ネットワークの識別子は、グラフ編集手段14において対象とする遺伝子ネットワークを構成する点・辺を選択してその連結成分IDを取得し、それからアニメーション表示手段111において現在選択されている遺伝子ネットワークの識別子をグラフ編集手段14の被選択点・辺メモリ504を介して取得することによって得ることが可能である。また、遺伝子ネットワークを表示する順序を示す番号は、アニメーション編集コマンドにより、再構成グラフメモリ5にあるグラフ再構成データDAに、アニメーション表示対象であるフラグと共に付与される。上述の操作を繰り返すことにより、アニメーションの表示対象となる再構成グラフを設定することができる。

【0131】次に、アニメーション表示コマンドについて説明する。このコマンドを処理するには、再構成グラフメモリ5に存在するグラフ再構成データDAにおいてアニメーション表示に関する設定が為されている必要がある。アニメーション表示コマンドがアニメーション表示手段111に入力されると、アニメーション表示対象となっているグラフ再構成データDBをレイアウトするようグラフ配置手段7にアニメーション表示コマンドを出力する。

【0132】図8に基づいて説明したようなレイヤー表示方法で、各表示順序に対応する再構成グラフ間で同じ要素が同じ位置に表示されるようにグラフ配置手段7により整列させる。また、ユーザの操作によってグラフ再構成データDAのレイアウトを編集する場合も考えられる。この場合は、グラフ配置手段7により整列処理を省くことが可能である。その場合は、グラフ編集手段14により配置グラフメモリ8に保存されているグラフ配置データDBの点・辺の位置や向きを変更する。アニメー-

ション表示手段111はこのグラフ配置データDBを全てのアニメーション表示対象について配置グラフメモリ8から取得し、アニメーション表示順序に対応する番号を付してアニメーションデータメモリ112に一旦格納する。

【0133】次に、アニメーションデータメモリ112に一旦格納されているグラフ配置データDBをアニメーション表示手段111を介してグラフ表示手段9に表示順に出力する。この処理により、グラフ表示手段9のグラフ配置データDBは表示メモリ10を介してディスプレイ装置6に送られ、配置グラフに対応する遺伝子ネットワーク図がディスプレイ装置6に表示される。

【0134】遺伝子ネットワーク研究支援システム1Cにアニメーション表示処理部100を設けたので、細胞内の分子機構を理解する上で極めて重要である、環境条件の変化によって変化する遺伝子ネットワークの変化を容易に把握できる。

【0135】遺伝子ネットワーク研究支援システム1、1A、1B、1Cは以上のように構成されており、遺伝子などの要素の情報とそれら要素間の関係の情報を蓄積したデータベースは予め与えられている。しかし、遺伝子に関する情報は日々更新され、新しい情報が常に増加している。最新の遺伝子ネットワーク研究にとって、これら最新の情報にアクセスすることは不可欠なことである。遺伝子に関する情報を蓄積したデータベースとしては、GENBANK、SWISS-PROTなどがある他、生物種を限定した詳細なデータベースや転写やスプライシングなどの特定の現象に着目したデータベースが構築、公開されている。予め決められた情報をこれらのデータベースから抽出し、抽出した情報を遺伝子情報データベース2に反映させるために、遺伝子ネットワーク研究支援システム1が常に最新の情報を遺伝子情報データベース2に反映するための情報抽出手段を有するようになることができる。このような目的を達成するための構成として図28に示すような情報抽出手段200を附加することもできる。

【0136】情報抽出手段200を説明すると、情報抽出手段200は図28に詳細に示されるように、抽出対象情報201と、遺伝子情報抽出手段202と、遺伝子情報保存手段203とから成っている。

【0137】抽出対象情報201にはどのデータベースのどこから情報を抽出するかを記述した抽出対象データが必要な抽出対象の個数分だけ格納されており、抽出対象データは遺伝子情報抽出手段202に送られる。遺伝子情報抽出手段202では、取得した抽出対象データを解釈し、インターネット等のネットワークを介してローカルあるいは遠隔地にあるデータベースにアクセスし、所望の情報を取得する。その後、取得した情報を遺伝子情報保存手段203に出力する。遺伝子情報保存手段203は、入力された情報を遺伝子情報データベース2に

(16) 2002-91991 (P2002-91991A)

出力して反映させるので、したがって、情報抽出手段200を設けることにより、遺伝子情報データベース2は常に最新の情報を有することができる。

【0138】遺伝子ネットワーク研究支援システム1は以上のような手段を付加することができる他、遺伝子ネットワーク研究支援システム1のグラフ再構成手段4においてシステムティックに再構成された遺伝子ネットワークに対して、ユーザが持っている実験結果や仮説をシステムの編集機能を使って遺伝子ネットワークに反映させるようにすることもできる。この機構により、ユーザが持っている仮説（例えば、なんらかの要素間の関係）を追加して再構成された遺伝子ネットワークと、遺伝子情報データベース2のみからシステムティックに再構成された遺伝子ネットワークとがどのように異なっているかが確認できるようになる。遺伝子ネットワーク研究支援システム1が備えている遺伝子情報データベース2には、過去に報告された情報、すなわち既知の情報が蓄積されており、この機構を使用することにより、ユーザが持っている仮説と既知の情報を遺伝子ネットワークという要素レベルよりも広い視点で比較することができるようになる。このような目的を達成するための構成として図29に示すような遺伝子情報追加機構300を付加することもできる。

【0139】遺伝子情報追加機構300を説明すると、遺伝子情報追加機構300は、遺伝子情報追加手段301と、追加遺伝子情報メモリ302とから成っている。

【0140】ユーザにより追加される1件以上の遺伝子情報から構成される遺伝子情報コマンドは、ユーザにより入力デバイス11から入力され、入力デバイス処理手段12を介して入力コマンド処理手段13に送られ、入力コマンド処理手段13において解釈された後、遺伝子情報追加手段301に出力される。遺伝子情報追加手段301は、入力された遺伝子情報から成る遺伝子情報追加コマンドを追加遺伝子情報メモリ302に出力すると共に、遺伝子情報データベース2と追加遺伝子情報メモリ302の両情報源を参照してグラフを再構成するための命令をグラフ再構成手段4に対して出力する。グラフ再構成手段4ではこの命令により遺伝子情報取得手段3を介して遺伝子情報データベース2と追加遺伝子情報メモリ302との2つの情報源を参照してグラフを再構成する。以後のグラフィカル表示のための処理は、図1に示した遺伝子ネットワーク研究支援システム1と同様である。

【0141】したがって、ディスプレイ装置6にグラフィカル表示された遺伝子ネットワークには遺伝子情報データベース2と遺伝子情報追加機構300とによりユーザが追加した遺伝子情報に由来する情報が含まれている。すなわち、これら2つの情報源に由来する関係情報が併記されることとなる。

【0142】上述の如く遺伝子情報追加機構300を付

加したので例えば以下のような効果を期待することができる。

【0143】まず最初に遺伝子情報データベース2に基づいた遺伝子ネットワークを再構成する。このとき再構成の条件・制約を指定しても良い。次に、再構成された遺伝子ネットワークに対して、ユーザが新たな関係情報を追加し、同じ条件・制約で遺伝子ネットワークを再構成することができる。この結果、着目しているネットワークに違いがあれば、その仮説の意味を理解する上で役に立つ。このような機能は、グラフ再構成の条件・制約が複雑（条件・制約の個数が多いなど）な場合などの理由により、再構成されるネットワークの予測が困難である場合に特に有効である。

【0144】上記と同様に、ユーザが追加した関係情報を遺伝子情報データベース2に基づき、遺伝子ネットワークを再構成し、この処理によって得られた遺伝子ネットワークの中に遺伝子情報データベース2に由来する関係とユーザによって追加された関係情報が含まれている場合、両者の違いを表示される色や形を異ならせることにより表示形態に反映させるようにしたので、ある関係が、遺伝子情報データベース2だけか、あるいはユーザによって追加された関係情報だけなのか、それとも両方の情報源に共通して存在するものなのかを容易に把握することができる。

【0145】

【発明の効果】本発明による効果は以下の通りである。

(1) 要素情報及び関係情報を蓄積したデータベースを構築しておき、ユーザの要求に応じて動的にそのデータベースを参照して遺伝子ネットワークを再構成して、ユーザがそれをグラフィカルな画像として確認したり、インクランクティブに編集できるようにしたので、データベースに蓄積された遺伝子に関する多種、多様、大量な実験データ、及び厖大な情報解析結果から遺伝子ネットワークを構築するのを支援することができる。

(2) 本発明によるシステムに特定の連結成分だけを指定するための手段を設けたので、所望の連結成分を選択して表示するようにすることができる。

(3) グラフ再構成手段に対して特定の連結成分だけを抽出するための条件・制約を指定することができる手段を設けたので、厖大な個数の連結成分から所望の連結成分だけをユーザが選択することができる。

(4) 要素間関係推論手段を含むように構成したので、グラフィカル表示された複雑な遺伝子ネットワークを単純化することができ、ユーザが複雑な遺伝子ネットワークを容易に理解できるようになる。

(5) 要素間関係推論手段の結果が表示されている場合に、要素間関係に対応するエッジを指定したり、指定されたエッジに対応する要素間関係に介在する関係情報に対応するエッジを表示したりするためのステップを設けたので、表示されている結果の推論のプロセスを視覚的

(17) 2002-91991 (P2002-91991A)

に理解することができる。

(6) 既存の文献あるいはデータベースから情報を抽出する情報抽出手段を設けたので、本発明によるデータベースに常に最新の情報を蓄積することができる。

(7) レイヤー表示手段を設けたので、複数のグラフを重ね合わせて表示することができ、それぞれのグラフの要素間の類似性や関連性などを容易に把握することができる。

(8) レイヤー表示機構において、複数レイヤー間で重複する又は重複しないノード、エッジを強調表示するようにしたので、遺伝子ネットワークが複雑な場合においても、複数レイヤー間のグラフの類似性及び相違部分を容易に把握することができる。

(9) ユーザから問合せパターンであるグラフとそれに関する付随情報についての条件とを入力するステップと、そのグラフに同型なグラフを有する再構成グラフを検索するステップと、検索結果に対して付随情報について照合を行うステップとを設けたので、類似するパターンを有するグラフを検索することができる。

(10) 本発明のデータベースに蓄積された要素情報および関係情報に付随する情報に応じて、それらに対応するノード、エッジの色の濃度を変更するようにしたので、ユーザーの信頼性の判定基準を反映した遺伝子ネットワークの表示を行うことができる。

(11) 付随する情報に応じて、ノード、エッジに対応する表示アリミティブの視覚的特徴を変更する手段を設けたので、表示された遺伝子ネットワーク図から直観的にノード、エッジに付随する情報を理解することができる。

(12) データベースに各要素に付随する情報を予め付随遺伝子情報として蓄積しておくと共に、ノードあるいはエッジを指定する手段を設けたので、指定されたノードあるいはエッジに対応する要素の付随遺伝子情報をデータベースから取得し、取得された付随遺伝子情報を表示することができ、要素や関係の意味を容易に理解することができる。

(13) 関係情報を追加、削除する手段を設けたので、ユーザーが持っている仮説を追加して再構成された遺伝子ネットワークと、本発明によるシステムのデータベースのみから再構成された遺伝子ネットワークとがどのように異なるかを確認することができる。

(14) 複数の遺伝子ネットワーク図を保存し、それを連続的に表示する機構を設けたので、例えば環境条件の変化に伴う遺伝子ネットワークの変化を容易に把握することができる。

【図面の簡単な説明】

【図1】本発明による遺伝子ネットワーク研究支援システムの実施の形態の一例を示すシステム構成図。

【図2】遺伝子データベース内に蓄積されている要素情報と関係情報との一例を示す模式図。

【図3】グラフ再構成手段によって実施されるグラフ再構成を説明するための図。

【図4】再構成されたグラフデータが遺伝子ネットワークとしてディスプレイ装置にグラフィカル表示される一例を示す図。

【図5】再構成グラフメモリに格納されているグラフデータを説明するための図。

【図6】配置グラフメモリに格納されているグラフ配置データを説明するための図。

【図7】本発明による遺伝子ネットワーク研究支援システムのグラフ編集手段の実施の形態の一例を示すシステム構成図。

【図8】本発明の第2の実施の形態を示すシステム構成図。

【図9】図8に示した連結成分選択部の詳細構成図。

【図10】連結成分選択部により抽出された連結成分がディスプレイ装置に表示されるリストの一例を示す図。

【図11】連結成分選択部において実施される連結成分選択の概念を説明するための説明図。

【図12】本発明の第2の実施の形態に設けられる条件・制約指定部の詳細構成図。

【図13】演繹推論部の演繹推論の実施による遺伝子ネットワークの単純化と新たな関係の生成との一例を説明するための説明図。

【図14】演繹推論部の類似関係の考慮による遺伝子ネットワークの単純化と新たな関係の生成との一例を説明するための説明図。

【図15】図8に示した演繹推論部の詳細構成図。

【図16】演繹推論部の演繹推論の繰り返し実施による遺伝子ネットワークの単純化と新たな関係の生成との一例を説明するための説明図。

【図17】演繹推論部による多段推論結果の表示の一例を示す図。

【図18】表示形態変更部により遺伝子ネットワークの点・辺の濃度の変更を説明するための説明図。

【図19】図8のシステムに設けられる表示形態変更部の詳細構成図。

【図20】レイヤー表示部により遺伝子ネットワーク図をレイヤー表示する実施の形態の一例を説明するための説明図。

【図21】レイヤー間の共通部分と差分を強調表示の一例を説明するための説明図。

【図22】遺伝子ネットワークの類似パターンを説明するための説明図。

【図23】図8に示した検索部の詳細構成図。

【図24】本発明の第3の実施の形態を示すシステム構成図。

【図25】図24に示した付随遺伝子情報表示手段により表示される再構成グラフの付随遺伝子情報の一例を説明するための模式図。

(18) 2002-91991 (P2002-91991A)

【図26】付随遺伝子情報表示手段によりグラフィカル表示された再構成グラフの付随遺伝子情報の表示の一例を示す図。

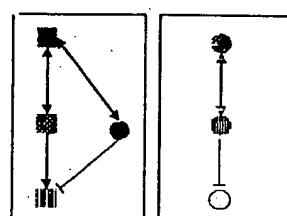
【図27】本発明の第4の実施の形態を示すシステム構成図。

【図28】図27に示すシステムに設けられる情報抽出手段の詳細構成図。

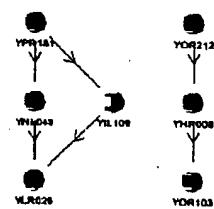
【図29】図27に示すシステムに設けられる遺伝子情報追加機構の詳細構成図。

【符号の説明】

1 遺伝子ネットワーク研究支援システム	52 表示形態条件メモリ
2 遺伝子情報データベース	60 レイヤー表示部
3 遺伝子情報取得手段	61 グラフ情報保存読取手段
4 グラフ再構成手段	62 グラフ情報
5 再構成グラフメモリ	63 レイヤー表示手段
6 ディスプレイ装置	70 検索部
7 グラフ配置手段	71 同型グラフ検索手段
8 配置グラフメモリ	72 問合せパターンメモリ
9 グラフ表示手段	73 検索結果メモリ
10 表示メモリ	81 付随遺伝子情報表示手段
11 入力デバイス	82 表示形態条件指定手段
12 入力デバイス処理手段	83 表示形態条件メモリ
13 入力コマンド処理手段	100 アニメーション表示処理部
14 グラフ編集手段	111 アニメーション表示手段
15 付随遺伝子情報表示手段	112 アニメーションデータメモリ
20 連結成分選択部	200 情報抽出手段
21 再構成グラフ選択指定手段	201 抽出対象情報
22 選択再構成グラフメモリ	202 遺伝子情報抽出手段
30 条件・制約指定部	203 遺伝子情報保存手段
31 再構成条件指定手段	300 遺伝子情報追加機構
32 再構成条件メモリ	301 遺伝子情報追加手段
40 演繹推論部	302 追加遺伝子情報メモリ
41 推論規則データベース	500 コマンド処理選択手段
42 推論規則選択手段	501 2次元位置算出部
43 推論規則メモリ	502 2次元位置メモリ
44 推論手段	503 点・辺配置情報取得手段
45 推論結果メモリ	504 被選択点・辺メモリ
50 表示形態変更部	511 点・辺選択コマンド処理部
51 表示形態条件指定手段	512 点・辺移動コマンド処理部

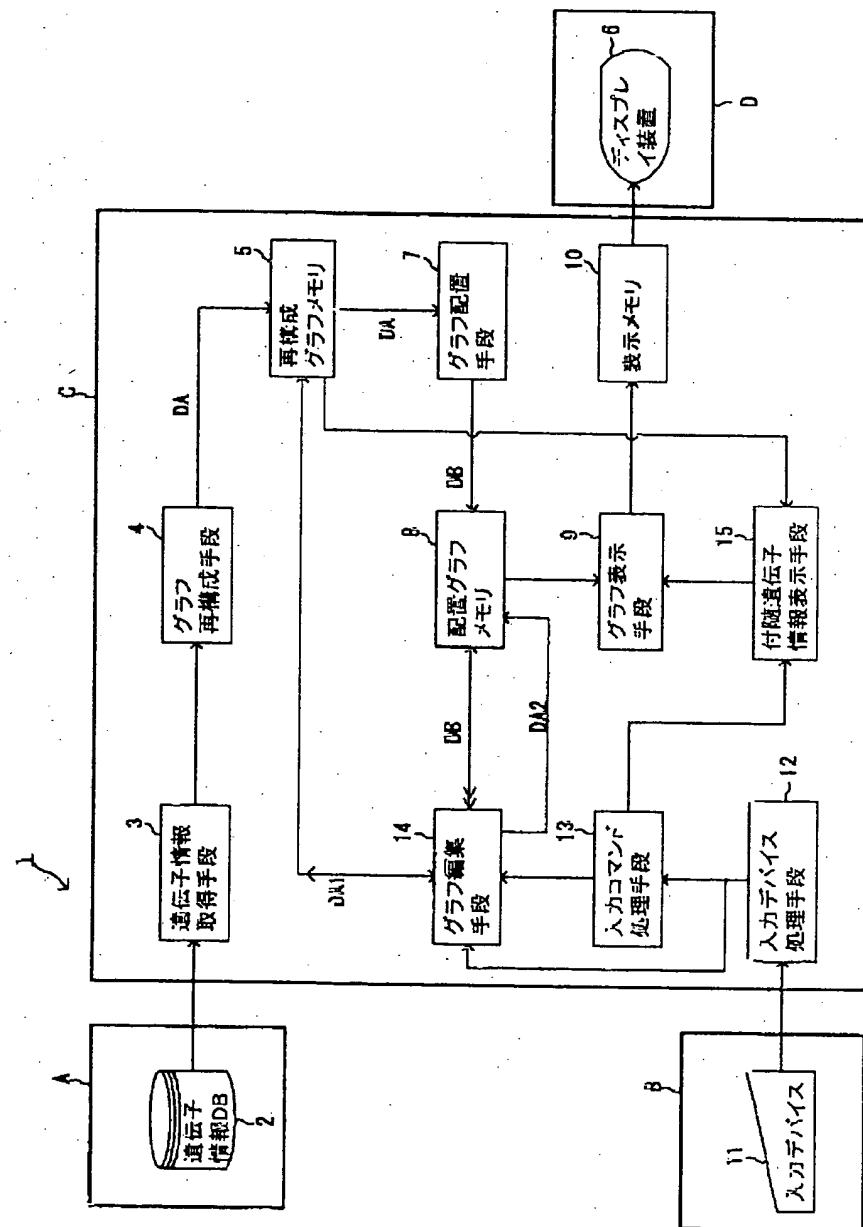


【図4】



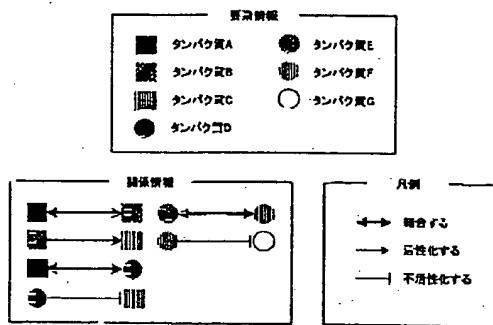
(19) 2002-91991 (P2002-91991A)

【図1】



(20) 2002-91991 (P2002-91991A)

【図2】



【図5】

要素説明

連結成分	要素ID	ラベル	部位	部位内局所	生物種	シノニム
1	00001	YNL011	Docking protein	cytoplasmic	yeast	SEC23, NOC1, P9705, 14
1	00002	YNL049	co-activator	cytoplasmic	yeast	SFB2, LSS1, YNL250S
1	00003	YLR026	Transcription factor	nuclear	yeast	SENL, L1742
1	00004	YIL109	co-activator	cytoplasmic	yeast	SEC24, AWU1
2	00006	YDR212	Docking protein	cytoplasmic	yeast	SITE4, BMD2, TOR50-2
2	00006	YHR005	Helicase	cytoplasmic	yeast	CPA1, BAC1, SEC1, CDC7
2	00007	YDR103	Transcription factor	nuclear	yeast	SITE4, BMD2, TOR50-2

結合情報

連結成分	要素ID	結合	結合	結合タイプ	位置	情報の由来
1	00001	YPR181	00002	YNL049	結合する	1.0 実験 36298
1	00002	YNL049	00003	YLR026	反応化する	0.7 増殖活性分析 20571
1	00003	YPR181	00004	YIL109	結合する	1.0 増殖 26298
1	00004	YIL109	00003	YLR026	不活性化する	0.5 増殖活性分析 20571
2	00006	YDR212	00006	YHR005	結合する	1.0 実験 36298
2	00006	YHR005	00007	YDR103	活性化する	0.6 実験 16349

【図6】

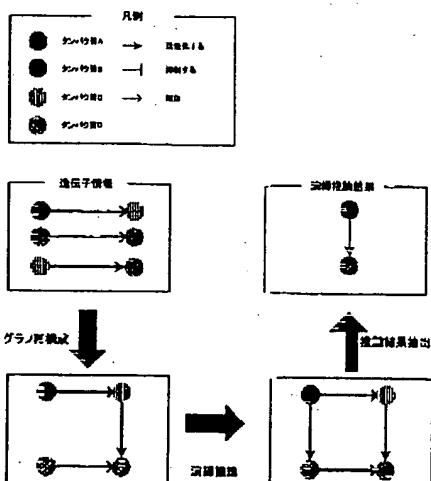
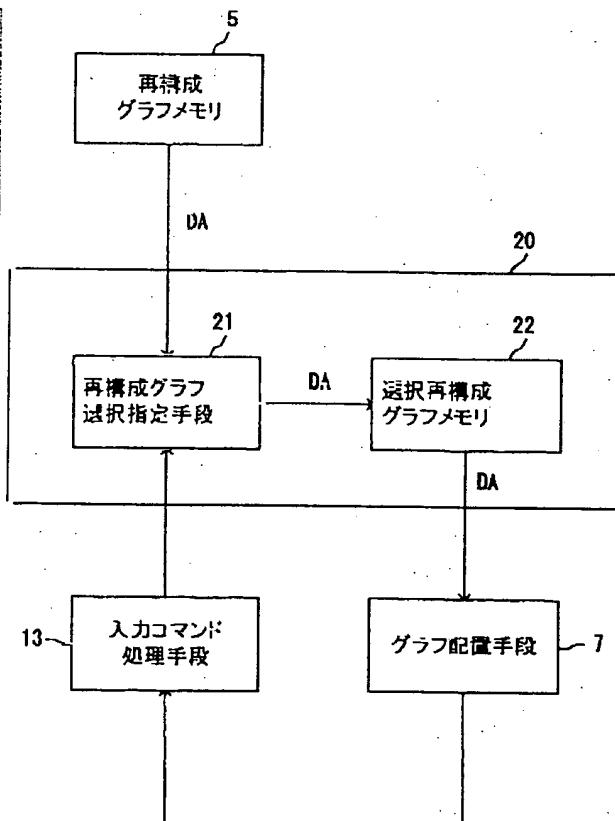
要素説明

連結成分	要素ID	座標	
		X	Y
1	00001	1.00	1.00
1	00002	1.00	3.00
1	00003	1.00	5.00
1	00004	3.00	3.00
2	00005	5.00	1.00
2	00006	5.00	3.00
2	00007	5.00	5.00

結合情報

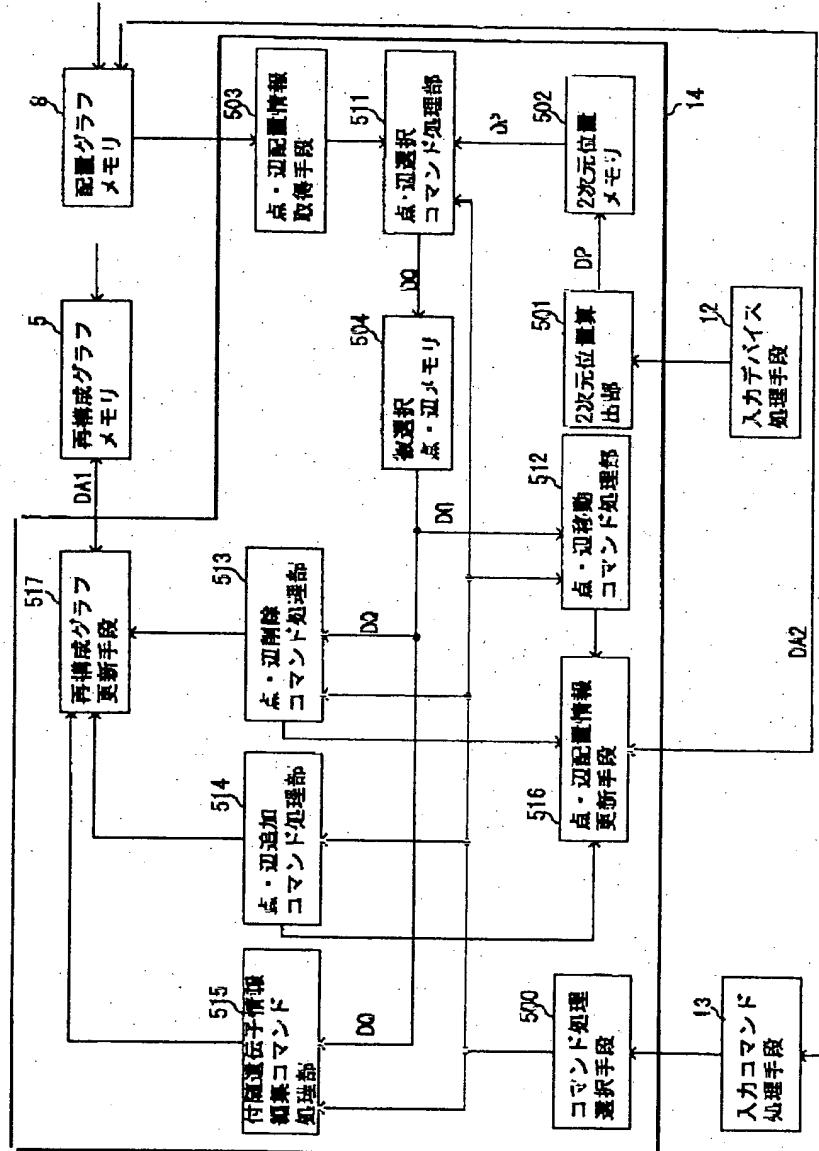
連結成分	要素ID	結合	結合座標		結合座標
			要素ID	要素ID	
1	00001	00002	1.00	1.00	1.00
1	00002	00003	1.00	3.00	3.00
1	00003	00004	1.00	5.00	5.00
1	00004	00005	3.00	3.00	3.00
2	00005	00006	5.00	1.00	1.00
2	00006	00007	5.00	3.00	3.00

【図14】

再構成
グラフメモリ

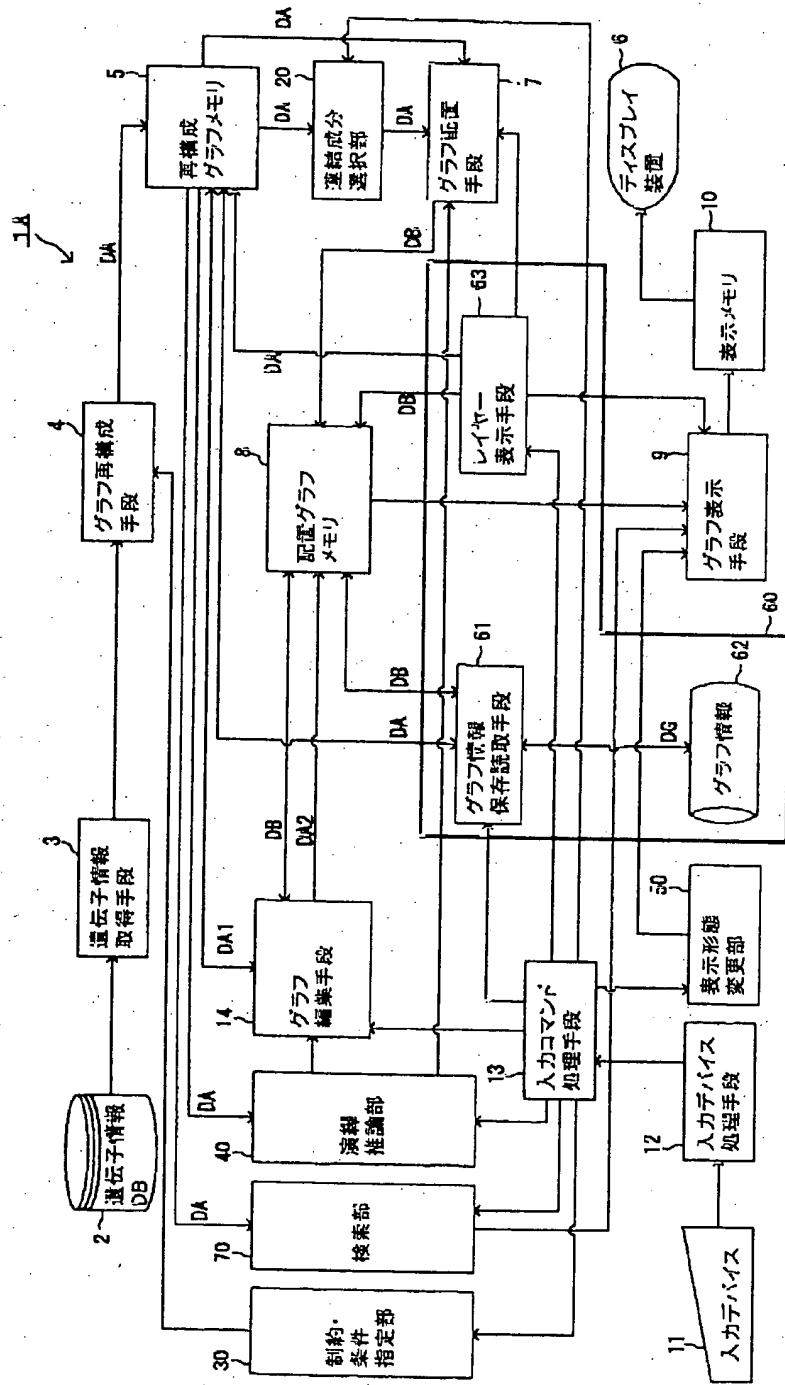
(21) 2002-91991 (P2002-91991A)

[図7]



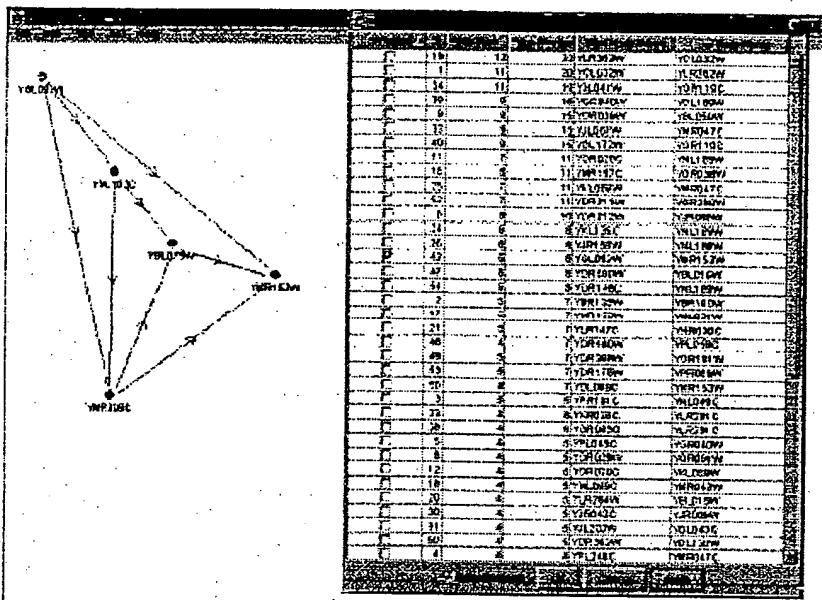
(22) 2002-91991 (P2002-91991A)

【図8】

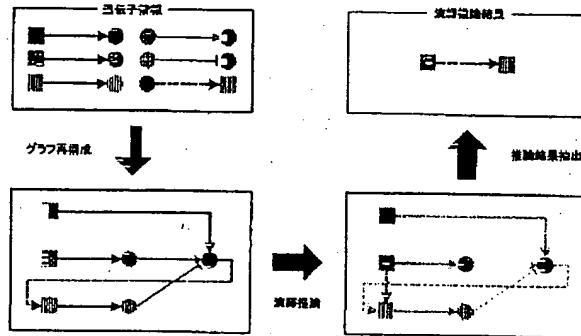
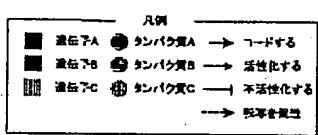


(23) 2002-91991 (P 2002-91991A)

【図10】



【図13】

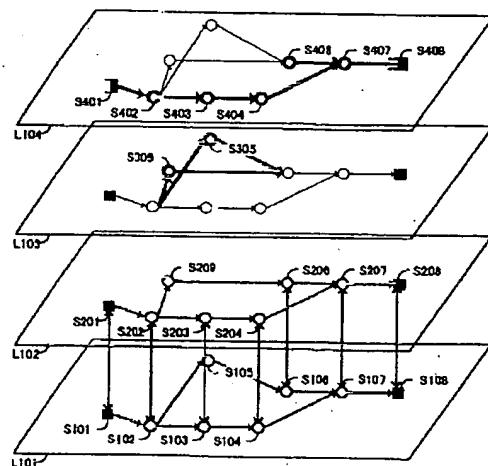


【図21】

■ 二重子
 ○ タンパク質
 → コードする
 → 活性化する
 ↔ 同一である
 — 強調線

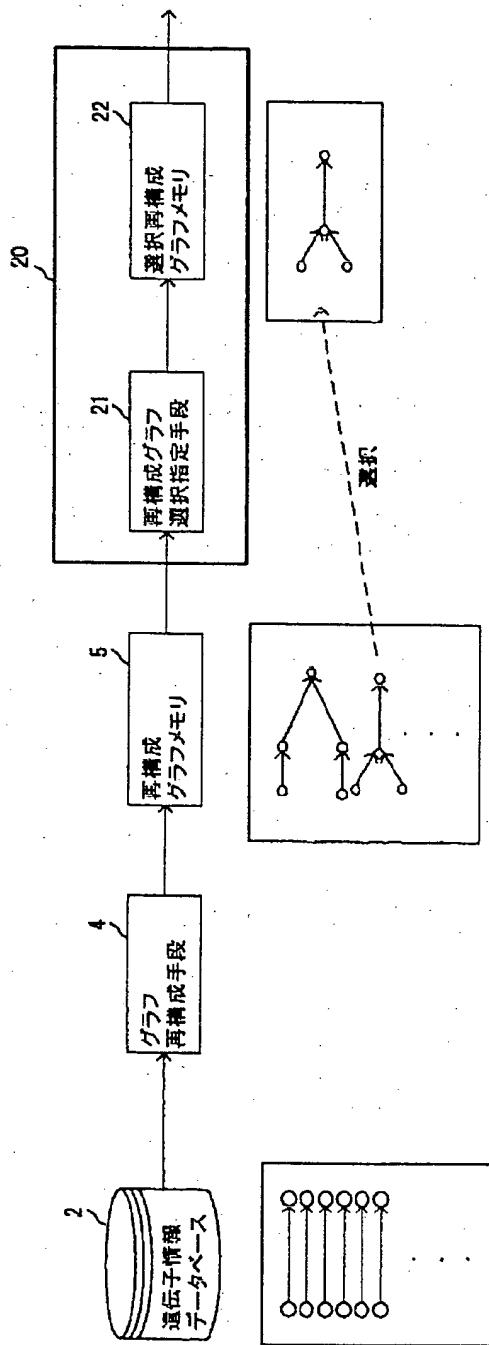
L101レイヤー-1(環境条件A)
 L102レイヤー-2(環境条件B)
 L103レイヤー-3(差分強調)
 L104レイヤー-4(差分強調)

S101～S108 実験(レイヤー-1)
 S201～S209 実験(レイヤー-2)
 S301,S309 実験(レイヤー-3)
 S401～S408 実験(レイヤー-4)



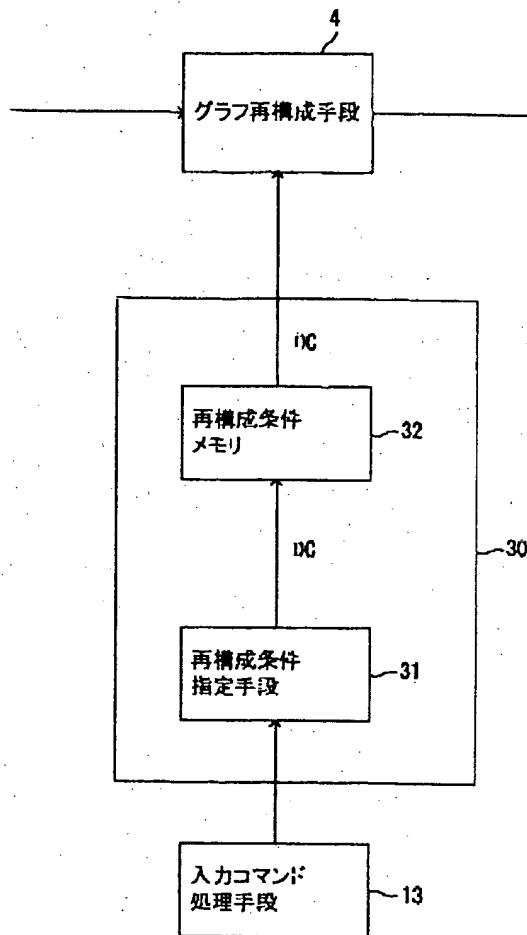
(24) 2002-91991 (P2002-91991A)

[図11]

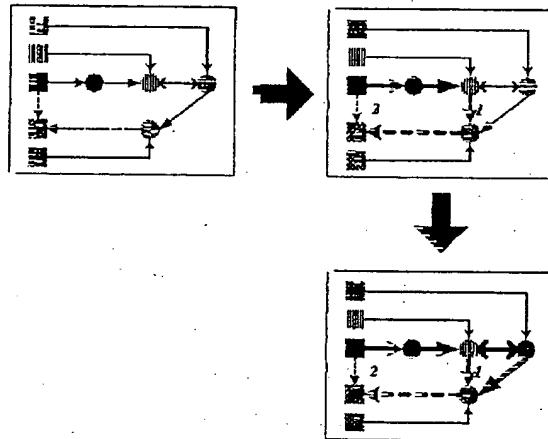


(25) 2002-91991 (P2002-91991A)

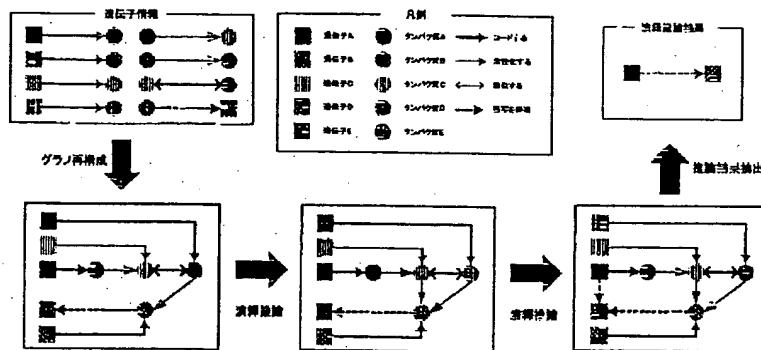
【図12】



[図17]

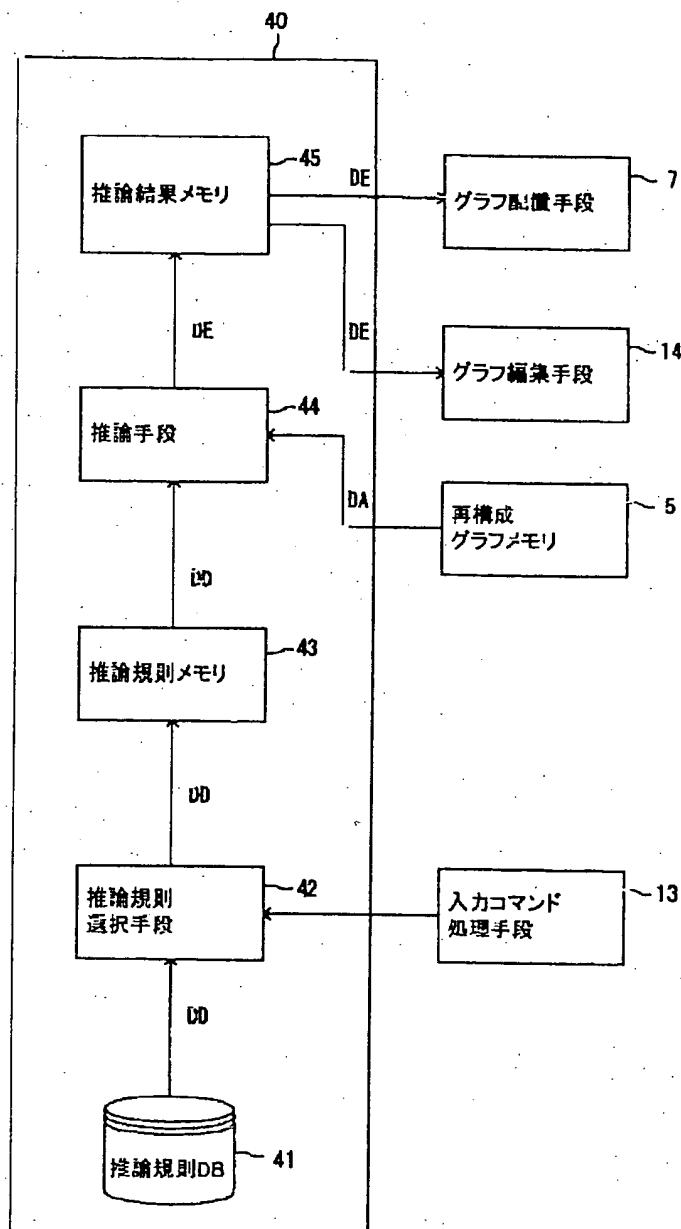


【图16】

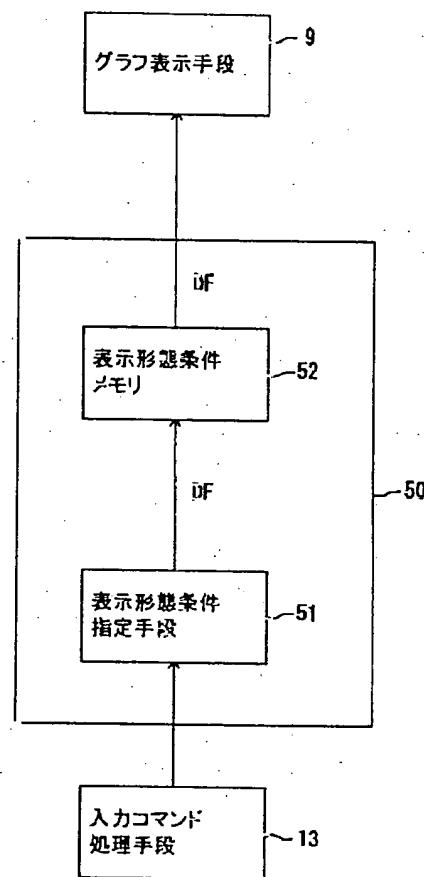


(26) 2002-91991 (P2002-91991A)

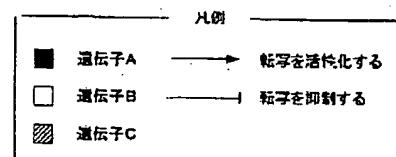
【図15】



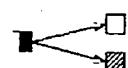
【図19】



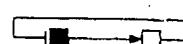
【図22】



例(1)

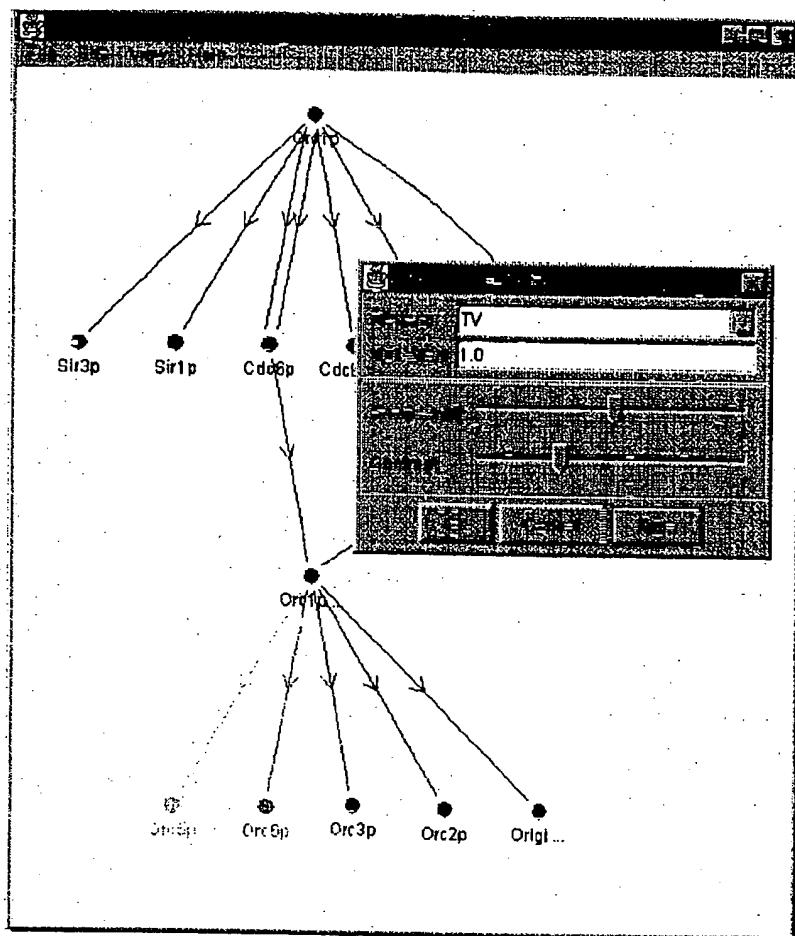


例(2)

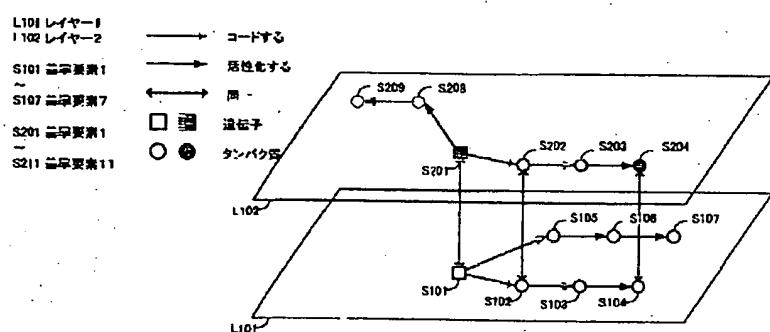


(27) 2002-91991 (P2002-91991A)

【図18】

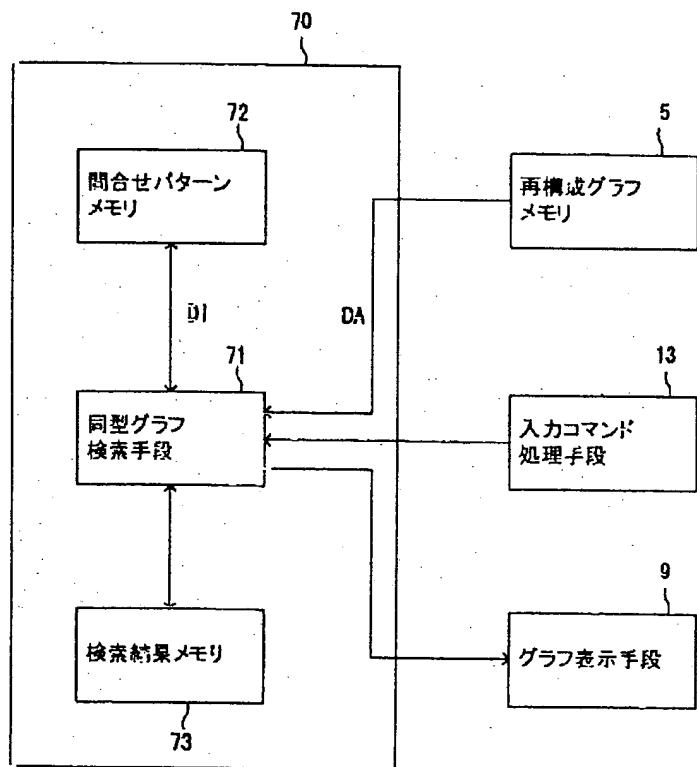


【図20】

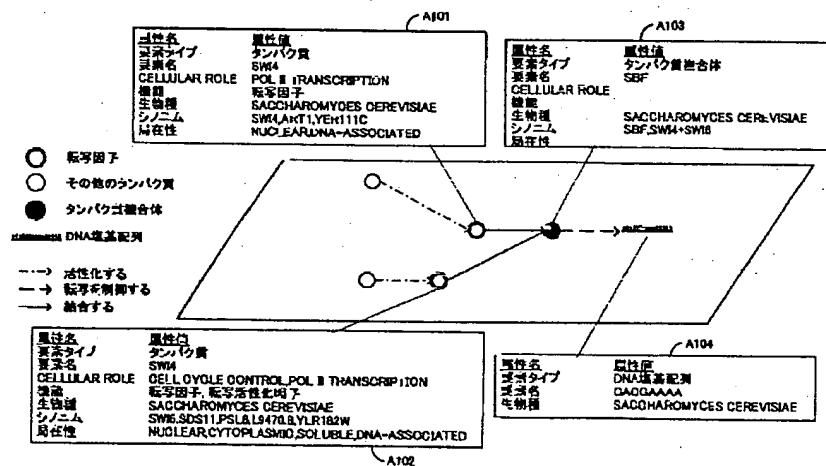


(28) 2002-91991 (P2002-91991A)

【図23】

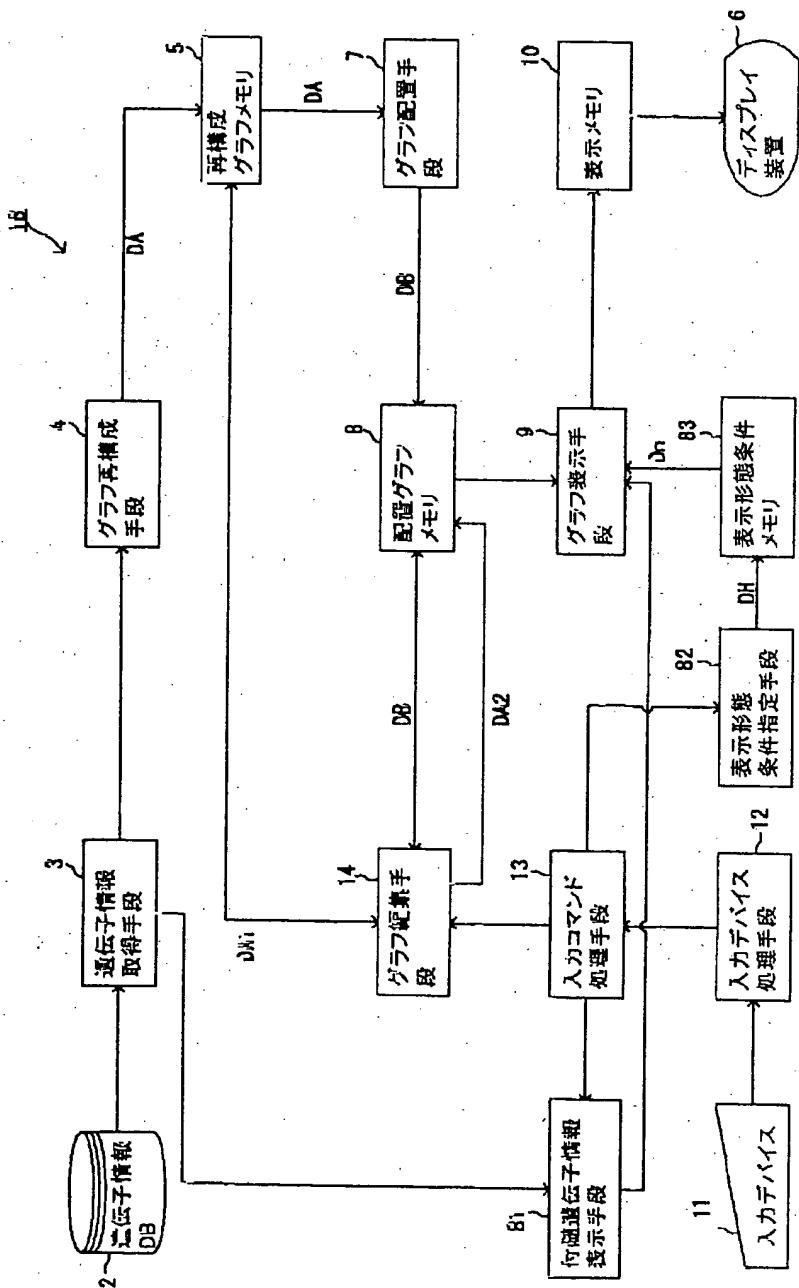


【図25】



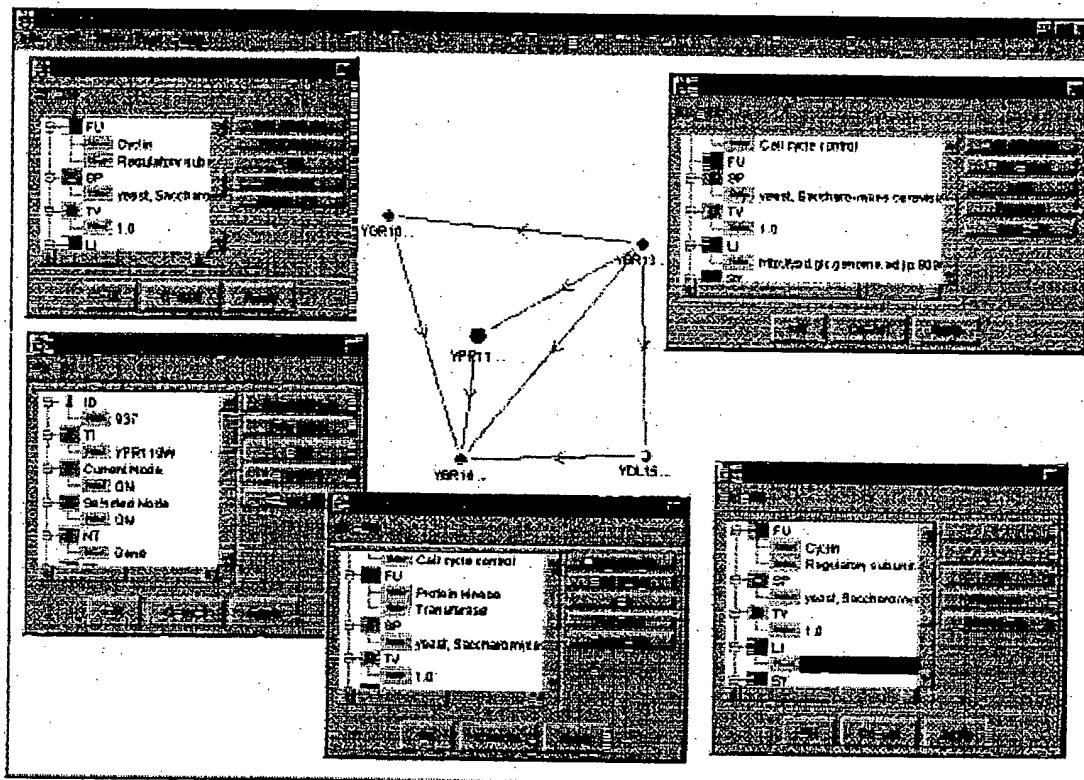
(29) 2002-91991 (P2002-91991A)

〔図24〕



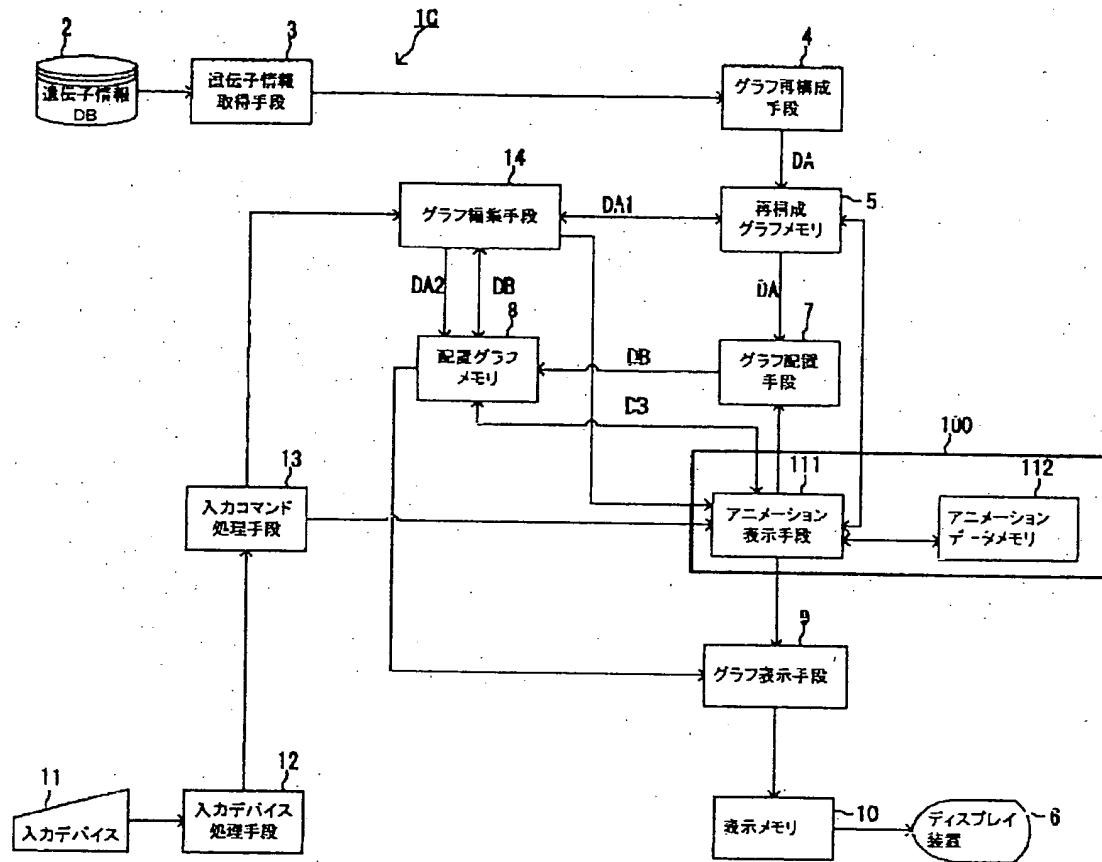
(30) 2002-91991 (P2002-91991A)

[図26]



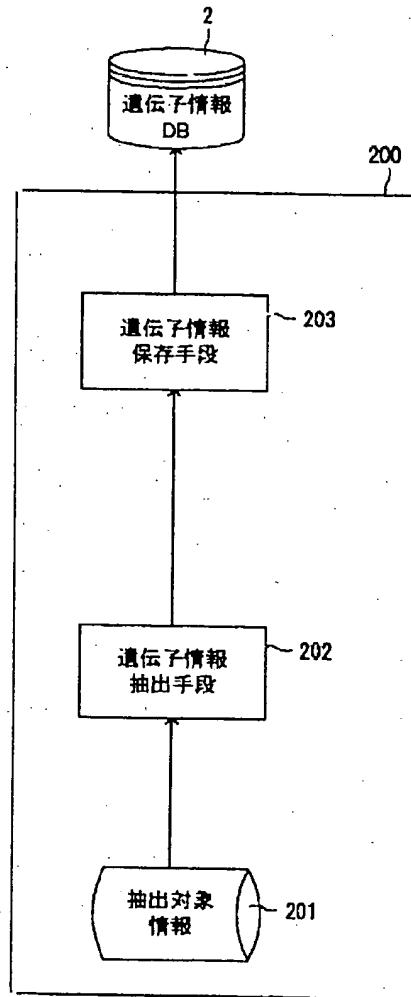
(31) 2002-91991 (P2002-91991A)

【図27】



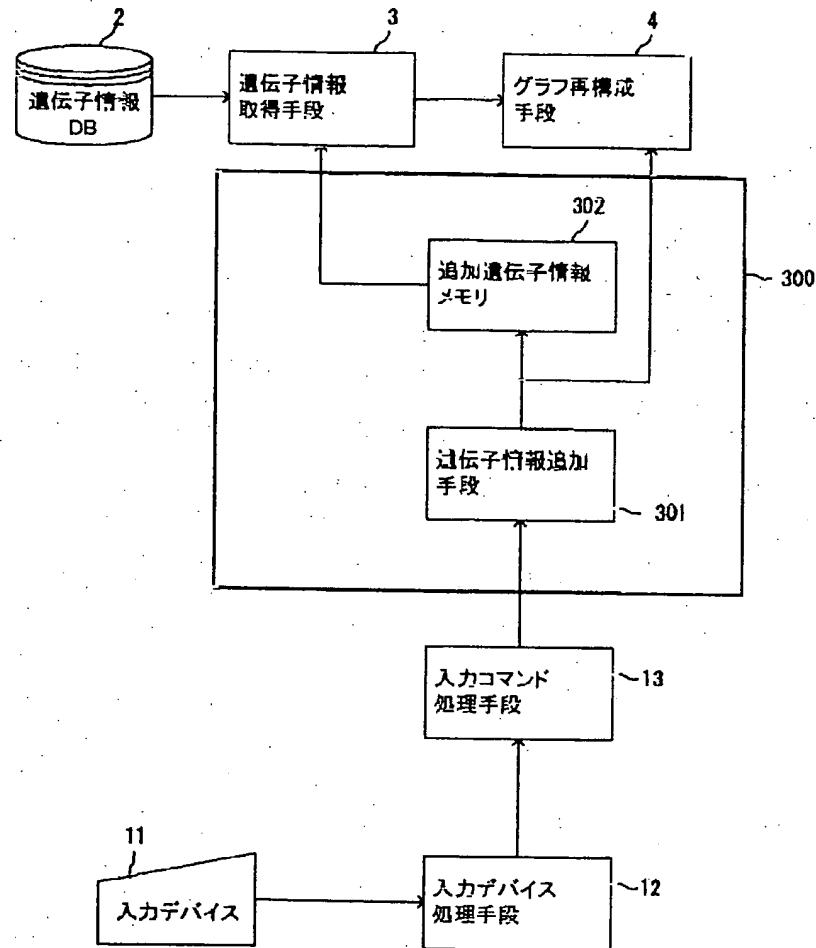
(32) 2002-91991 (P 2002-91991A)

【図28】



(33) 2002-91991 (P2002-91991A)

【図29】



フロントページの続き

(72)発明者 深川 浩志

東京都江東区新砂一丁目3番3号 インテック・ウェブ・アンド・ゲノム・インフォマティクス株式会社内

(72)発明者 柴垣 ゆかり

富山県富山市下新町3番23号 インテック・ウェブ・アンド・ゲノム・インフォマティクス株式会社内

(72)発明者 島野 英明

富山県富山市下新町3番23号 インテック・ウェブ・アンド・ゲノム・インフォマティクス株式会社内

(72)発明者 高木 利久

東京都江東区越中島1-3-16-609

(72)発明者 伊藤 隆司

石川県金沢市平和町3-20-10 平和宿舎
(C) 54-31

F ターム(参考) 4B024 AA19 AA20 CA01 CA11 HA20

4B029 AA23 BB20

5B075 ND02 PQ02 PQ14 PQ22 UU19

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- BLACK BORDERS**
- IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- FADED TEXT OR DRAWING**
- BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- SKEWED/SLANTED IMAGES**
- COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- GRAY SCALE DOCUMENTS**
- LINES OR MARKS ON ORIGINAL DOCUMENT**
- REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.